

CDS

89..1897
/gene="BCHHE"
/codon_start=-1
/product="butyrylcholinesterase precursor"
/protein_id="AAC06261.1"
/db_xref="pid:92981241"
/db_xref="GI:2981241"
/translation="MOSKGTITISOFILRLFLMLVIGKSHREDIITTKNKVRGM
NLPVLDTVAFLGIPYADPPLGRLEKPKKQPTLRKSDIMNAIKYNSCYONADQSEF
GFGSEMMNPNTDLSBDCLYLNWIPTRPKRNTVMWLYGGEGFTGSLPYDGKE
LARVERIVYKSNYRVGALGFALPGNPEVPGMGLEDDOLALQWQKNIALFGNPK
SVTLFGESAGASVSLHLSPROSPFTTRAILOSGSNPAVMASLDEAKNRLITLAK
FIGCSKENDREIITKLRNKPDEIILNELLYVSDTLISVNEGPPVDDGFLDMPRTL
LQGGOKRTOIIVYKNDKDEGATILYVAGSGESKNDNSITTRKREFGRLKTYPPGVSEF
GREAILFTYDLDLDORAEKRYREALDVLADNITICPALEFTTKSELSGNNAFTYFE
HRSQSLPWPWGVHGVETEFVFGLEPLERKRVYTAEEILTSISINWANKRGINDP
NCTONNSTRWPAFSTDKYTLNASESPKRYTLRAQOCRFWTLFFPKVLEMTGNIDE
AERERAGFRRNNYMDMKQFNQDYTSKESCAGL"
173..1894
/gene="BCHHE"
/note="serum cholinesterase; pseudocholinesterase; serine
esterase"
/product="butyrylcholinesterase"

BASE COUNT 598 a 375 c 409 g 550 t

ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 1932;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1397 TTCTCAGAA 1405
|||||

Cp 33 ttctcagaa 25

RESULT 106
LOCUS CCTPMY13 1972 bp DNA VRT 07-MAR-1996
DEFINITION Quail gene for alpha-tropomyosin, exon 13.
ACCESSION X16241.X14883
NID 962672
VERSION X16241.1 GI:62672
KEYWORDS alpha-tropomyosin; alternative splicing; tropomyosin.
SOURCE quail.
ORGANISM Coturnix coturnix
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
REFERENCE 1 (bases 1 to 1972)
AUTHORS Lindquester,G.J.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1989) to the EMBL/GenBank/DBJ databases
REFERENCE 2 (bases 1 to 1972)
AUTHORS Lindquester,G.J., Flach,J.E., Fleenor,D.E., Hickman,K.H. and
Devlin,R.B.
TITLE Avian tropomyosin gene expression
JOURNAL Nucleic Acids Res. 17 (5): 2099-2118 (1989)
MEDLINE 89183613
REFERENCE 3 (bases 1 to 1972)
AUTHORS Kristensen,T., Lopez,R. and Prydz,H.
TITLE An estimate of the sequencing error frequency in the DNA sequence
databases
JOURNAL DNA Seq. 2 (6), 343-346 (1992)
MEDLINE 93075997
REMARK Erratum: [[published erratum appears in DNA Seq 1993;3(5):337]]
see X16230 - X16241 for alpha-tropomyosin gene exons 1-13. splicing
pathways:
skeletal muscle type : exons 1/3/5-10/11/12
smooth muscle type : exons 1/2/5-10/13
nonmuscle type : exons 4/5-10/7
Data kindly reviewed (18-SEP-1989) by Lindquester G.J.
FEATURES
source 1..1972
/organism="Coturnix coturnix"
/strain="Japanese."

misc_feature
/db_xref="taxon:9091"
/clone_lib="lambda Charon4A"
251..1180
/label=ex13
/usedin=X16230:trop_mrna
251..333
/label=stop
/usedin=X16230:tropsm.cds
/usedin=X16230:trop.cds
972..977
polyA_signal
polyA_signal
1145..1150
polyA_site
1180
misc_feature
1923..1972
/note="putative VECTOR contamination M13mp18"
/citation={3}

BASE COUNT 561 a 372 c 421 g 617 t 1 others

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 1972;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 188 TTCTCAGAA 196
|||||

Cp 25 ttctcagaa 33

RESULT 107
LOCUS NIB00060 1974 bp RNA INV 03-SEP-1997
DEFINITION Naegleria italica 18S ribosomal RNA gene, partial sequence.
ACCESSION U80060
NID 92351572
VERSION U80060.1 GI:2351572
KEYWORDS
SOURCE Naegleria italica.
ORGANISM Naegleria italica
Eukaryota; Mitochondrial eukaryotes; Heterolobosoa;
Schizopyrenida; Vahlkampfiidae; Naegleria.
REFERENCE 1 (bases 1 to 1974)
AUTHORS Einvik,C., Decatur,W.A., Embley,T.M., Vogt,V.M. and Johansen,S.
TITLE Naegleria nucleolar introns contain two group I ribozymes with
different functions in RNA splicing and processing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1974)
AUTHORS Embley,T.M. and Dyal,P.L.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1996) Zoology, Natural History Museum, Cromwell
Road, London SW7 5BD, UK
FEATURES
source 1..1974
/organism="Naegleria italica"
/strain="AB-T-F3"
/db_xref="taxon:36334"
RNA
BASE COUNT 507 a 414 c 524 g 528 t 1 others

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 1974;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1217 TTCTCAGAA 1225
|||||

Cp 33 ttctcagaa 25

RESULT 108
LOCUS AF084366 1979 bp DNA BCT 23-AUG-1998
DEFINITION Lactobacillus bulgaricus glutamyl-tRNA synthetase gene, complete
cds.
ACCESSION AF084366
NID 93450894

NID 910126
 VERSION X03076.1 GI:10126
 KEYWORDS actin; inverted repeat.
 SOURCE Strongylocentrotus franciscanus.
 ORGANISM Strongylocentrotus franciscanus
 Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 1876)
 FORAN, D.R., JOHNSON, P.J. and MOORE, G.P.
 Evolution of two actin genes in the sea urchin *Strongylocentrotus franciscanus*
 JOURNAL J. Mol. Evol. 22 (2), 108-116 (1985)
 MEDLINE 86062634
 COMMENT Introns 1 (pos. 547-770) is composed of inverted repeats. Data kindly reviewed (05-FEB-1986) by D.R. Foran.

FEATURES
 source
 Location/Qualifiers
 1..1876
 /organism="Strongylocentrotus franciscanus"
 /db_xref="taxon:7665"
 1..180
 /number=1
 /join(181..546,771..1017,1209..1726)
 /codon_start=1
 /product="actin"
 /protein_id="CAA26878.1"
 /db_xref="PID:9295763"
 /db_xref="GI:295763"
 /db_xref="SWISS-PROT:P10991"
 /translation="MCDQVVALVINDSGNKKVAGFAGDAPRAVFPISVIGRPHQGV
 MVGKQKDSYVDEAGSRGILTKYPIEHGIVNWDMEKIMHTFYNELRVAPBEH
 PVLTLEPLNKANREKTOIMEFNSPAYVLAIOVALSYASGRTGYFEDGCV
 SHATVPEGVALPHAILRLDAGRDLDYMKILLTGYSFTTAAREIVDIEKIC
 YVALDFEDQGTAASSSSLEKSYELPDQGVTTTINERFRAEALFOAFLOMESAGH
 ETCYNISMKCDVDIRKDIYANTVLSGGSTMPGADRNOKKITLALAPPTMKITLAPP
 ERKYSVWIGSGISLASLSTFOOMWISKQYEDSGPSIYHRKCF"
 547..770
 /number=1
 /join(771..1017
 /number=2
 1018..1208
 /number=2
 1018..1876
 1209..1876
 /number=3
 504 a 466 c 384 g 522 t

BASE COUNT
 ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 1876;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1843 TTCTCAGAA 1851
 |||||
 Cp 33 ttctcagaa 25

RESULT 104
 LOCUS ABU09869 1877 bp DNA BCF 30-NOV-1995
 DEFINITION Azospirillum brasilense GTP cyclohydrolase II gene, complete cds.
 ACCESSION U09869
 NID 9497270
 VERSION U09869.1 GI:497270
 KEYWORDS
 SOURCE Azospirillum brasilense.
 ORGANISM Azospirillum brasilense
 Eubacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae; Azospirillum.

REFERENCE 1 (bases 1 to 1877)
 VAN BASTELAERE, E., KELLERS, V. and VANDERLEYDEN, J.
 Cloning and sequencing of the putative *Azospirillum brasilense* gene encoding GTP cyclohydrolase II

JOURNAL Gene 153 (1), 141-142 (1995)
 MEDLINE 95189096
 REFERENCE 2 (bases 1 to 1877)
 AUTHORS Van Bastelaere, E.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAY-1994) Els Van Bastelaere, F.A. Janssens
 Laboratory of Genetics, Applied Plant Sciences, Willem de Croylaan
 42, Heverlee, B-3001, Belgium

FEATURES
 source
 Location/Qualifiers
 1..1877
 /organism="Azospirillum brasilense"
 /strain="Sp7"
 /db_xref="taxon:192"
 /clone="pFAJ0103"
 355..1512
 /codon_start=1
 /transl_table=11
 /function="riboflavin biosynthesis"
 /product="GTP cyclohydrolase II"
 /protein_id="AA82170.1"
 /db_xref="PID:9497271"
 /db_xref="GI:497271"
 /translation="MYADAPSDSAPPRAVLPNDAAARAVDPATALRCEAVAIETA
 DSSVGAIVSVESVAIDVORLVLTGAPVLAIVTRRAYLTKMGESTGVVALSLPFC
 LTADEAALADPEHRPDGMDPLTATAMPDSRETAADVLARLARLPAIVAIPATD
 HTGSAEMAELHDLIVRARDIADYRVHVVTRVAEAVPLSGAENTSIAFRPID
 GGEPEHLAIIVGNVAVGEPVLAHLSECFDGLAGLDCDQGLGAIARHSGV
 LLYLAQDRGIGLVNKLRAVRIODRGFTVDANELLGFEDAEVYLPAAEMLQIGFT.
 AYRLMTNPKLQIARCGIEVERPHIFPAHGHNGIYRTAERGHNF"
 1548..1566
 stem_loop 280 a 674 c 650 g 273 t

BASE COUNT
 ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1507 TTCTCAGAA 1515
 |||||
 Oy 25 ttctcagaa 33

RESULT 105
 LOCUS AF053483 1932 bp mRNA NAM 23-MAR-1998
 DEFINITION *Felis catus* butyrylcholinesterase precursor (BCHE) mRNA, complete cds.
 ACCESSION AF053483
 NID 92981240
 VERSION AF053483.1 GI:2981240
 KEYWORDS
 SOURCE cat.
 ORGANISM *Felis catus*
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Carnivora; Fissipedia; Felidae; Felis.

REFERENCE 1 (bases 1 to 1932)
 BARTELS, C.F., XIE, W.-H., MILLER-LINDHOLM, A.K. and LOCKRIDGE, O.
 Sequence and characterization of domestic cat acetylcholinesterase
 and butyrylcholinesterase
 Unpublished

JOURNAL 2 (bases 1 to 1932)
 REFERENCE Bartels, C.F., Xie, W.-H., Miller-Lindholm, A.K. and Lockridge, O.
 AUTHORS Direct Submission
 TITLE Submitted (11-MAR-1998) Eppley Institute, University of Nebraska
 JOURNAL Medical Center, 600 S. 42nd St., Omaha, NE 68198-6805, USA

FEATURES
 source
 Location/Qualifiers
 1..1932
 /organism="Felis catus"
 /db_xref="taxon:9685"
 /tissue_type="pituitary"
 1..1932
 /gene="BCHE"
 89..172
 /gene="BCHE"

gene
 sig_peptide

CDS

/note="lamin B receptor"
/evidence=experimental
103.2016
/codon_start=1
/product="lamin B receptor"
/protein_id="CA68758.1"
/db_xref="PID:962934"
/db_xref="GI:62934"
/db_xref="SWISS-PROT:P23913"
/translation="MPNKKYADGEVYMKRMGSGVLYVEQVTSYDDASHLYTVYKRG
TELLAKESDILQSFQKRSQSSSSPSRSRSRSRSGPRAKGRSSSHREH
KEDKKKIIOETSLAPKPSENNTRRYNCEPSTERNDSKLEQOKIKPVEHERV
DOYLSRREKKKEEYIAEKIFEATITPERPSSKTELEFGGFGFMFLPAT
VLVIMCKODDPSLMNPPLPALESIMETKVPFVFLMEFQALFTYLPYIGKVEGL
PLSNPKIQYINQFYAFLLTAALIGITLYQFELHTLYDFYQPAVSAASMAISI
VLYTRSLAPEEDLAPGNSGIVLYDFETGHELNPRIGSFQKTELPOLIGMVAI
NLAMLEMKIHNSMPSLSMILVNSFQLAYVALNMEEVLTMTDITHGFGFMFL
EGDLVWPEVYSIQAFYLVGHPIAISMPVAAITILNIGIYIFRSANSONNRRNP
ADPKLSYKVIPTATGKGLVYTGMMGFRRHNPYSGLDIIIMALMSLPCGFNHLIFYFY
IFYICLVHREARDEHHKKKYGILAMERYCQRVPYTHISLHLESHSTYLICKLYTSH
LCTWSVCTLGKFKH"

BASE COUNT 585 a 442 c 481 g 615 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2123;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1596 TTCACGAA 1604
|||||||

Cp 33 ttctcagaa 25

RESULT 124
LOCUS A03759 2136 bp DNA PAT 27-NOV-1993
DEFINITION H. sapiens (pX153) gene for serum albumin (reverse complement).
ACCESSION A03759
NID 9488986
VERSION A03759.1 GI:488986
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2136)
AUTHORS Latta,M., Mayaux,J.F. and Sarmientos,P.
TITLE Method for the synthesis of mature human serum albumin
JOURNAL Patent: EP 0236210-A 6 09-SEP-1987;
GENETICA

FEATURES
SOURCE 1. 2136
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 567 a 459 c 436 g 674 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 2136;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 468 TTCACGAA 476
|||||||

Cp 33 ttctcagaa 25

RESULT 125
LOCUS S73803 2147 bp mRNA INV 01-MAR-1995
DEFINITION HLC-32-32 kda hyaline layer component [sea urchins, embryos, mRNA,
2147 nt].
ACCESSION S73803
NID 9688099
VERSION S73803.1 GI:688099
KEYWORDS

SOURCE sea urchins embryos.
ORGANISM Echinoidea
Eukaryota; mitochondrial eukaryotes; Metazoa; Echinodermata;
Echinozoa.

REFERENCE 1 (bases 1 to 2147)
AUTHORS Brennan,C. and Robinson,J.J.
TITLE Cloning and characterization of HLC-32, a 32-kDa protein component
of the sea urchin extraembryonic matrix, the hyaline layer
Dev. Biol. 165 (2), 556-565 (1994)

JOURNAL MEDLINE
REMARK 95046875
GenBank staff at the National Library of Medicine created this
entry [NCBI g15bq 156426] from the original journal article.
This sequence comes from Fig. 3.

FEATURES
SOURCE 1. 2147
Location/Qualifiers
/organism="Echinoidea"
/db_xref="taxon:7625"
871. 1998
/gene="HLC-32"
871. 1998
CDS
/gene="HLC-32"
/note="32 kda hyaline layer component; This sequence comes
from Fig. 3."
/codon_start=1
/product="HLC-32"
/protein_id="AA832327.1"
/db_xref="PID:9688100"
/db_xref="GI:688100"
/translation="MLPWCFGSDPTTRGIGCLKASISAKFSCPSRRSGSVGVALIN
PMVLSPAEQNOQNAISAKFEFLRMFPADSNVSNRSDNVDVIGRLALIRENA
NIAEQVDLRQVLRSDLSVTLCMPSDKAVQKROELPAKLRDQEMKNLVAMVVS
GRVSSQISDNQKQVSLNGAKLRNRYARADSKKMTIANGARVADBRKASGILHYD
KVYPLPVGVMETLADNPAPFMSVVDLKOAGLEDEVNDSPTIVLPTAAQOALPA
GYDDDKKDPAGKRLNLYKHVYSDVYKVSLSGGRASGODEISVANDSDQILL
NKQSDQSKASVILRDIPTTNGYVQVDRILVRSQKHFVLY"

BASE COUNT 578 a 539 c 481 g 549 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 2147;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 586 TTCACGAA 594
|||||||

Cp 33 ttctcagaa 25

RESULT 126
LOCUS ECU6825 2157 bp DNA INV 19-OCT-1998
DEFINITION Encephalitozoon cuniculi gene encoding trehalose 6-phosphate
phosphatase.
ACCESSION A1006825
NID 93776048
VERSION A1006825.1 GI:3776048
KEYWORDS trehalose 6-phosphate phosphatase.
SOURCE Encephalitozoon cuniculi.
ORGANISM Encephalitozoon cuniculi
Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.

REFERENCE 1 (bases 1 to 2157)
AUTHORS Duffieux,F.
TITLE Direct Submision
JOURNAL Submitted (08-JUN-1998) Duffieux F., Universite B. Pascal-UPESA
CNRS 6023, Protistologie Mol. et Cell. des Parasites Opportunistes,
Campus des Cezeaux, 24 av. des Landais, Aubiere, 63177, FRANCE

REFERENCE 2 (bases 1 to 2157)
AUTHORS Duffieux,F., Peyret,P., Roe,B.A. and Viveses,C.P.
TITLE Putative trehalose 6-phosphate phosphatase from the chromosome I of
Encephalitozoon cuniculi (Microspora)

JOURNAL Unpublished
FEATURES
SOURCE 1. 2157
Location/Qualifiers
/organism="Encephalitozoon cuniculi"
/isolate="mouse"

cds

/db_xref="taxon:6035"
/chromosome="1"
/dev_stage="spore"
/lab_host="MDCK cells"
1..2157
/EC_number="3.1.3.12"
/codon_start=1
/product="trehalase 6-phosphate phosphatase"
/protein_id="CA07262.1"
/db_xref="PID:e1316154"
/db_xref="PID:g3776049"
/db_xref="GI:3776049"
/translation="MKIIIVAAEELPICASRMPESEAIKNNISTPLKKIPHPDKASKIQ
IEFGDRKESHLPAKPHFVPMDDDEHFEVAPGPSPSEEDDCRIEESMKA
YRCPIYOKYKYSKMEILLGKNTPEFNOENPHAMFARYKRYENRMKEIMEYE
EGDVVWMDSLFLPGMLGNSIPVGSASAPSSLLKCIPIFNEQITSSILCCYIEF
NESSKESFDLVSOQKTFGMDPGYDINEPLTCVKGKIDKVDLKMSEVHEEG
LGKGVLLPSDSOTHLGVEAYLSRGKETVLETRVGLNSDSQAVMLREYL
EINKVLSRETPASDEFTSMKRCDCLEPADVCSLEGIYVYNNYDEFDIAD
EINENLMORGESGESSEVYIGKMEKKEKFTSLISGMEYDVLPEKPRISLS
MDOTGKKVDAKKRPGIRKKRKEEAEVEITINKEANARIVNDEKSKARTLVND
YDGTITIVARPPAPPTOEIKDLILILGKICRYVETGSEVDCDFPKETIEVFE
HGACHRIDGKMKEGTIPQKDLAWRIQDFLARYGSELEKKTGIAHFRNPSPLG
VKQARALFELMRVKDKYVKGNHVIIEVRSKSCAEKIEEGFVLCAGDVADDEMF
DVCKGYTIKVGDOSTSAAYRVKDPENFRMLGRULE"

BASE COUNT 621 a 416 c 607 g 513 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 2157;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 932 TTCGAGAA 940
|||||||
25 ttctgagaa 33

RESULT 127
LOCUS CHRCYP4502 2160 bp mRNA VRT 23-AUG-1994
DEFINITION Chicken phenobarbital-inducible cytochrome P450 mRNA, complete cds.
ACCESSION M25469
M25469.1 GI:532117
VERSION cytochrome P450.
KEYWORDS Gallus gallus cDNA to mRNA.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
1 (bases 1 to 2160)
Hansen, A.J. and May, B.K.
Sequence of a chicken phenobarbital-inducible cytochrome P450 cDNA:
Regulation of two P450 mRNAs transcribed from different genes
DNA 8, 179-191 (1989)
JOURNAL 89251079
MEDLINE On Aug 24, 1994 this sequence version replaced gi:341348.
COMMENT Location/Qualifiers
FEATURES
source
1..2160
/organism="Gallus gallus"
/db_xref="taxon:9031"
40..1515
/note="phenobarbital-inducible"
/codon_start=1
/product="cytochrome P450"
/protein_id="AA48743.1"
/db_xref="PID:g532118"
/db_xref="GI:532118"

BASE COUNT 753 a 362 c 423 g 638 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 2176;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1395 TTCGAGAA 1403
|||||||
25 ttctgagaa 33

RESULT 128
LOCUS SSAGAPD 2176 bp DNA BCT 17-SEP-1997
DEFINITION S.solfataricus agapd gene.
ACCESSION x66409
x66409.1 GI:510674
VERSION x66409.1 GI:510674
KEYWORDS glyceraldehyde-3-phosphate dehydrogenase.
SOURCE Sulfolobus solfataricus
ORGANISM Archaea; Crenarchaeota; Sulfolobales; Sulfolobus.
1 (bases 1 to 2176)
Arcati, P.
Direct Submission
Submitted (18-MAY-1992) P. Arcati, Dipartimento di Biochimica e,
Biotechnologie Mediche, University of Naples, Via S Pansini 5,
I-80131 Napoli, ITALY
2 (bases 1 to 2176)
Arcati, P., Russo, A.D., Iannicello, G., Gallo, M. and Bocchini, V.
Nucleotide sequence and molecular evolution of the gene coding for
glyceraldehyde-3-phosphate dehydrogenase in the thermophilic
archaeobacterium Sulfolobus solfataricus
Biochem. Genet. 31 (5-6), 241-251 (1993)
94082761
Location/Qualifiers
1..2176
/organism="Sulfolobus solfataricus"
/strain="WT-4 (ATCC 49255)"
/db_xref="taxon:2287"
311..1336
/gene="agapd"
311..1336
/gene="agapd"
/EC_number="1.2.1.12"
/codon_start=1
/transl_table=11
/product="glyceraldehyde 3-phosphate dehydrogenase"
/protein_id="CA447040.1"
/db_xref="PID:g809766"
/db_xref="GI:809766"
/db_xref="SWISS-PROT:P39460"
/translation="MINVANNGTIGTGKRVADAIKIQPMKLVGAKTSNPEAFIAH
RNGYRIYALPOOSIKRFEESGIPVAGTIVYLIKTSIDVDTTPNGVAGQKPYIDQ
RNAIFOGGEKAEVADISFSAICNVNPAKSKYIRVYSCNTTALLRICIVNKYKVEN
VRATIVBRADOKEVKGPINSLVDPAPVSHADVNSVIRNLDAIVAPPTL
MHHFNITIKDKVEKKDILSVLENPRVILISSKDAEATLVARLAKRDRDI
PEVMTISDSIYVADDEVMLRAYHOSIYVPEPIDAIRSMKLSMSEDSKRTINESLG
ILNGYLI"

BASE COUNT 753 a 362 c 423 g 638 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 2176;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1395 TTCGAGAA 1403
|||||||
25 ttctgagaa 33

RESULT 128
LOCUS SSAGAPD 2176 bp DNA BCT 17-SEP-1997
DEFINITION S.solfataricus agapd gene.
ACCESSION x66409
x66409.1 GI:510674
VERSION x66409.1 GI:510674
KEYWORDS glyceraldehyde-3-phosphate dehydrogenase.
SOURCE Sulfolobus solfataricus
ORGANISM Archaea; Crenarchaeota; Sulfolobales; Sulfolobus.
1 (bases 1 to 2176)
Arcati, P.
Direct Submission
Submitted (18-MAY-1992) P. Arcati, Dipartimento di Biochimica e,
Biotechnologie Mediche, University of Naples, Via S Pansini 5,
I-80131 Napoli, ITALY
2 (bases 1 to 2176)
Arcati, P., Russo, A.D., Iannicello, G., Gallo, M. and Bocchini, V.
Nucleotide sequence and molecular evolution of the gene coding for
glyceraldehyde-3-phosphate dehydrogenase in the thermophilic
archaeobacterium Sulfolobus solfataricus
Biochem. Genet. 31 (5-6), 241-251 (1993)
94082761
Location/Qualifiers
1..2176
/organism="Sulfolobus solfataricus"
/strain="WT-4 (ATCC 49255)"
/db_xref="taxon:2287"
311..1336
/gene="agapd"
311..1336
/gene="agapd"
/EC_number="1.2.1.12"
/codon_start=1
/transl_table=11
/product="glyceraldehyde 3-phosphate dehydrogenase"
/protein_id="CA447040.1"
/db_xref="PID:g809766"
/db_xref="GI:809766"
/db_xref="SWISS-PROT:P39460"
/translation="MINVANNGTIGTGKRVADAIKIQPMKLVGAKTSNPEAFIAH
RNGYRIYALPOOSIKRFEESGIPVAGTIVYLIKTSIDVDTTPNGVAGQKPYIDQ
RNAIFOGGEKAEVADISFSAICNVNPAKSKYIRVYSCNTTALLRICIVNKYKVEN
VRATIVBRADOKEVKGPINSLVDPAPVSHADVNSVIRNLDAIVAPPTL
MHHFNITIKDKVEKKDILSVLENPRVILISSKDAEATLVARLAKRDRDI
PEVMTISDSIYVADDEVMLRAYHOSIYVPEPIDAIRSMKLSMSEDSKRTINESLG
ILNGYLI"

BASE COUNT 753 a 362 c 423 g 638 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 2176;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1395 TTCGAGAA 1403
|||||||
25 ttctgagaa 33

polya_signal

BASE COUNT 619 a 505 c 470 g 566 t

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2160;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1823 TTCGAGAA 1831
|||||||
25 ttctgagaa 33

RESULT 128
LOCUS SSAGAPD 2176 bp DNA BCT 17-SEP-1997
DEFINITION S.solfataricus agapd gene.
ACCESSION x66409
x66409.1 GI:510674
VERSION x66409.1 GI:510674
KEYWORDS glyceraldehyde-3-phosphate dehydrogenase.
SOURCE Sulfolobus solfataricus
ORGANISM Archaea; Crenarchaeota; Sulfolobales; Sulfolobus.
1 (bases 1 to 2176)
Arcati, P.
Direct Submission
Submitted (18-MAY-1992) P. Arcati, Dipartimento di Biochimica e,
Biotechnologie Mediche, University of Naples, Via S Pansini 5,
I-80131 Napoli, ITALY
2 (bases 1 to 2176)
Arcati, P., Russo, A.D., Iannicello, G., Gallo, M. and Bocchini, V.
Nucleotide sequence and molecular evolution of the gene coding for
glyceraldehyde-3-phosphate dehydrogenase in the thermophilic
archaeobacterium Sulfolobus solfataricus
Biochem. Genet. 31 (5-6), 241-251 (1993)
94082761
Location/Qualifiers
1..2176
/organism="Sulfolobus solfataricus"
/strain="WT-4 (ATCC 49255)"
/db_xref="taxon:2287"
311..1336
/gene="agapd"
311..1336
/gene="agapd"
/EC_number="1.2.1.12"
/codon_start=1
/transl_table=11
/product="glyceraldehyde 3-phosphate dehydrogenase"
/protein_id="CA447040.1"
/db_xref="PID:g809766"
/db_xref="GI:809766"
/db_xref="SWISS-PROT:P39460"
/translation="MINVANNGTIGTGKRVADAIKIQPMKLVGAKTSNPEAFIAH
RNGYRIYALPOOSIKRFEESGIPVAGTIVYLIKTSIDVDTTPNGVAGQKPYIDQ
RNAIFOGGEKAEVADISFSAICNVNPAKSKYIRVYSCNTTALLRICIVNKYKVEN
VRATIVBRADOKEVKGPINSLVDPAPVSHADVNSVIRNLDAIVAPPTL
MHHFNITIKDKVEKKDILSVLENPRVILISSKDAEATLVARLAKRDRDI
PEVMTISDSIYVADDEVMLRAYHOSIYVPEPIDAIRSMKLSMSEDSKRTINESLG
ILNGYLI"

BASE COUNT 753 a 362 c 423 g 638 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 2176;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1395 TTCGAGAA 1403
|||||||
25 ttctgagaa 33

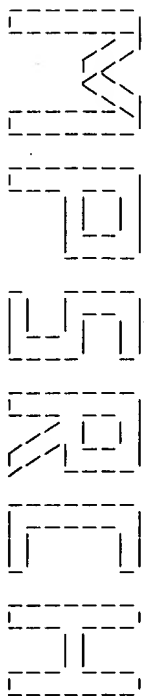
VERSION L23800.1 GI:438906
KEYWORDS
SOURCE Australian echidna.
ORGANISM Tachygllossus aculeatus
Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Monotremata; Tachygllossidae; Tachygllossus.
REFERENCE 1 (bases 1 to 2308)
AUTHORS Lee, M.-H., Shroff, R. and Hope, R.
TITLE Molecular characterization, mapping and evolution of a beta-globin gene from the Australian echidna, Tachygllossus aculeatus (Monotremata)
JOURNAL Unpublished (1994)
FEATURES
Source Location/Qualifiers
1..2308
/organism="Tachygllossus aculeatus"
/db_xref="taxon:9261"
/dev_stage="adult"
/sex="female"
/tissue_type="kidney"
91..105
/rpt_type="dispersed"
/rpt_unit="91..93"
join(<427..518,665..887,1353..>1481)
/gene="HBB"
/note="HBB"
/note="putative"
/number=1
join(427..518,665..887,1353..1481)
/gene="HBB"
/note="adult form; putative"
/codon_start=1
/product="beta-globin homolog"
/protein_id="BAB42177.1"
/db_xref="PID:9438907"
/db_xref="GI:438907"
/translation="MVLHSGSEKTAVTINMGHVNNELGEGALGRLVYVPMTORFFE
SEGLSDADAVNAKVGAKGAVLTSFGDLKNDNLKGTFAKLSEHCKRLHYDPE
NENRLGNLVVLAHNFSEKTPPEQAAMOKLVSGVSHALAHKYH"
519..664
/note="putative"
/number=1
665..887
/gene="HBB"
/note="putative"
/number=2
888..1352
/gene="HBB"
/note="putative"
/number=2
1353..>1481
/gene="HBB"
/note="putative"
/number=3
repeat_region
2120..2170
/rpt_type="dispersed"
BASE COUNT 542 a 577 c 674 g 515 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 2308;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 444 TTCTGAGAA 452
|||||
OY 25 ttctgagaa 33
RESULT 136 BTASPMATA 2347 bp mRNA MAM 08-NOV-1995
LOCUS B.taurus mitochondrial aspartate aminotransferase mRNA, complete
DEFINITION
CDS
ACCESSION Z25466

NID 9415323
VERSION Z25466.1 GI:415323
KEYWORDS aspartate aminotransferase.
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 2347)
AUTHORS Palmisano, A., Aurilia, V., Ferraro, L., Cubellis, M.V., Sanna, G. and Martino, G.
TITLE Nucleotide sequence of a cDNA coding for bovine mitochondrial aspartate aminotransferase
JOURNAL Int. J. Biochem. Cell Biol. 27 (5), 507-511 (1995)
MEDLINE 95368517
REFERENCE 2 (bases 1 to 2347)
AUTHORS Palmisano, A.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1993) Antonietta Palmisano, c/o Dip. Chimica Organica e Biologica, IABAM CNR, Via Mezzocannone, 16, Napoli, 80134, ITALIA
REMARK revised by [3]
REFERENCE 3 (bases 1 to 2347)
AUTHORS Palmisano, A.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1993) Antonietta Palmisano, c/o Dip. Chimica Organica e Biologica, IABAM CNR, Via Mezzocannone, 16, Napoli, 80134, ITALIA
COMMENT On Nov 11, 1993 this sequence version replaced gi:396282.
FEATURES
Source Location/Qualifiers
1..2347
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="clone p6B"
/dev_stage="adult"
/tissue_type="heart"
28..87
/EC_number="2.6.1.1"
/codon_start=1
/product="aspartate aminotransferase"
/protein_id="CA80960.1"
/db_xref="PID:9415324"
/db_xref="GI:415324"
/db_xref="SWISS-PROT:P12344"
/translation="MALHSGRFLSGVAAAFHPGLAAASARASGMVAHYEMGPDPPI
LGVEARPRDNRKKNLGVGAYRDDGKPYVPSVKRAEQAIAKRLDKDEYLPDGL
AERKASAEALGKNNELVLSGRVTVOTISGTALRIGASFLORFKSRDFTLPAP
TWGNHTPIFRDAGMQLQSYRKYDPTGDFGTGAIEDISKIPASVILLRACANPFG
VDPREPQMKEMATVVKNNLFAFDMAVQGSAGDNKDMVAHVFLEQIINCLQGS
YAKNMGKGERVGAFTVVCDAEAKVESQKILIRPNYSNPINIGARIASITLSP
DLRKMVHEVKGMDRIISMRTQVSNLKKKESGNHNOHIIIDQGMNCGYGLKREQVE
RLTKREPSIVMKGRISVAGVTSQNVAYLAHAIHQVYK"
BASE COUNT 585 a 607 c 579 g 576 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 2347;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1471 TTCTGAGAA 1479
|||||
CP 33 ttctcagaa 25
RESULT 137 SUSCIITIA 2350 bp DNA INV 15-JUN-1993
LOCUS Strongylocentrotus purpuratus T22 CyIIa actin gene regulatory domain.
DEFINITION
ACCESSION M64573
NID 9161460
VERSION M64573.1 GI:161460
KEYWORDS actin.
SOURCE S.purpuratus DNA.

```

/citation-[1]
/bound_molety-"P4"
/evidence-experimental
protein_bind
1963..1986
/citation-[1]
/bound_molety-"P5"
/evidence-experimental
protein_bind
2080..2104
/citation-[1]
/bound_molety-"P3A"
/evidence-experimental
protein_bind
2119..2135
/citation-[1]
/bound_molety-"P3B"
/evidence-experimental
protein_bind
2202..2235
/citation-[1]
/bound_molety-"P2I"
/evidence-experimental
protein_bind
2237..2264
/citation-[1]
/bound_molety-"P2II"
/evidence-experimental
protein_bind
2265..2274
/citation-[1]
/bound_molety-"P1"
/evidence-experimental
protein_bind
2308..2321
/citation-[1]
/bound_molety-"P1"
/evidence-experimental
BASE COUNT      699 a      527 c      432 g      692 t
ORIGIN
Query Match      100.0%; Score 9; DB 21; Length 2350;
Best Local Similarity 100.0%; Pred.No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db              9 TTCCTCAGAA 17
                |||||
CP              33 ttctcagaa 25

RESULT 138
LOCUS      ECGFPYRE      2375 bp      DNA      BCT      12-SEP-1993
DEFINITION E. coli DNA with 2 unidentified reading frames upstream the pyre gene.
ACCESSION  X14235
NID        942720
VERSION    GI:42720
KEYWORDS   pyre gene; unidentified reading frame.
SOURCE     Escherichia coli.
ORGANISM   Escherichia coli.
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
REFERENCE   1 (bases 1 to 2375)
AUTHORS    Poulsen, P.
TITLE      Direct Submision
JOURNAL    Submitted (27-JAN-1989) Poulsen P., Institute of Biological Chemistry K, University of Copenhagen, Solvgade 83, DK 1307 Copenhagen K, Denmark
REFERENCE   2 (bases 681 to 780)
AUTHORS    Andersen, J.T., Jensen, K.F. and Poulsen, P.
TITLE      Role of transcription pausing in the control of the pyre attenuator in Escherichia coli
JOURNAL    Mol. Microbiol. 5 (2), 327-333 (1991)
MEDLINE    91251761
FEATURES
             Location/Qualifiers
             1..2375
                /organism="Escherichia coli"
                /strain="K12"
                /db_xref="taxon:.562"
                /map="81.7 min"
```



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Oct 28 00:19:21 1999; Maspar time 88.65 seconds

281.403 Million cell updates/sec

Tabular output not generated.

Title: >US-08-963-288-1
(25-33) from US08963288.seq

Perfect Score: 9

N.A. Sequence: 25 tctctgaa 33

Comp: aagactctt

Scoring table: TABLE default

Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 646147 seqs, 1385953633 bases x 2

Post-processing: Minimum Match 100%

Listing first 200 summaries

Database:

emb158
1:em_ba1 2:em_ba2 3:em_fun 4:em_htg 5:em_hum1 6:em_hum2
7:em_in 8:em_om 9:em_ov 10:em_ov 11:em_pat 12:em_ph
13:em_pl 14:em_ro 15:em_sts 16:em_v1
genbank111
17:gb_ba1 18:gb_ba2 19:gb_htg1 20:gb_htg2 21:gb_in1
22:gb_in2 23:gb_om 24:gb_ov 25:gb_pat 26:gb_ph 27:gb_pl1
28:gb_pl2 29:gb_pl1 30:gb_pl2 31:gb_pl3 32:gb_ro
33:gb_st 34:gb_sts 35:gb_sy 36:gb_un 37:gb_v1

Statistics: Mean 5.304; Variance 1.626; scale 3.263

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	9	100.0	26 25	I46684	Sequence 663 from pate	8.17e+03
2	9	100.0	26 25	I46688	Sequence 667 from pate	8.17e+03
3	9	100.0	132 25	I32126	Sequence 12 from paten	8.17e+03
4	9	100.0	192 25	I84379	Sequence 37 from paten	8.17e+03
5	9	100.0	198 25	A59593	Sequence 6 from Patent	8.17e+03
6	9	100.0	242 23	BOVCO3	Bovine gene for conglu	8.17e+03
7	9	100.0	246 17	BBU65795	Bovine gene for conglu	8.17e+03
8	9	100.0	250 24	CYIRS	Carp (C. carpio) repeti	8.17e+03
9	9	100.0	256 24	APA011313	Acrossochellus paradox	8.17e+03
10	9	100.0	303 11	E01842	Genomic RNA of hop stu	8.17e+03
11	9	100.0	309 23	CNPMTB	Conenepus leuconotus m	8.17e+03
12	9	100.0	331 23	REFECYTB1	Rhinolophus ferrumequi	8.17e+03
13	9	100.0	367 17	RAU33923	Rickettsia australis c	8.17e+03

14	9	100.0	374 23	OAU47049	Ovis aries TG dinucleo	8.17e+03
15	9	100.0	383 17	RAU41752	Rickettsia akari citra	8.17e+03
16	9	100.0	385 21	EIR93401	E. invadens RNA gene (8.17e+03
17	9	100.0	396 17	UNSRRI096	Unknown organism, part	8.17e+03
18	9	100.0	399 21	MSU75307	Manduca sexta cellular	8.17e+03
19	9	100.0	402 23	AF044660	Rhinolophus hipposider	8.17e+03
20	9	100.0	436 18	AF114520	Uncultured gamma prote	8.17e+03
21	9	100.0	443 22	AF063883	Drosophila teissleri c	8.17e+03
22	9	100.0	470 23	SHPMICB	Ovis aries DNA microa	8.17e+03
23	9	100.0	479 21	LD004669	Lymantiria disper clone	8.17e+03
24	9	100.0	491 24	CHRCWB	Gallus domesticus DNA	8.17e+03
25	9	100.0	496 23	BOVCYTB04	Bovine cytochrome b-5	8.17e+03
26	9	100.0	536 24	HEF180	Heterodontus francisci	8.17e+03
27	9	100.0	578 17	BS16SRMU	B. subtilis (trpE26 mut	8.17e+03
28	9	100.0	583 25	I07689	Sequence 12 from Paten	8.17e+03
29	9	100.0	585 25	A48440	Sequence 5 from Patent	8.17e+03
30	9	100.0	662 25	I74318	Sequence 16 from paten	8.17e+03
31	9	100.0	668 24	AB023802S2	Gallus gallus Rd gene	8.17e+03
32	9	100.0	701 21	SCMGUPR	Schistosoma mansoni g1	8.17e+03
33	9	100.0	738 17	MSGBCGA85B	Mycobacterium bovis ge	8.17e+03
34	9	100.0	768 21	TAU32225	Theileria annulata cyt	8.17e+03
35	9	100.0	789 23	CHU34232	Capra hircus interfero	8.17e+03
36	9	100.0	808 17	BS16SRM5	B. subtilis 16S RNA ge	8.17e+03
37	9	100.0	855 27	AFJ3315	Aspergillus fumigatus	8.17e+03
38	9	100.0	890 25	A25392	Oligonucleotide.	8.17e+03
39	9	100.0	900 24	CHKRFBP	Chicken riboflavin-bin	8.17e+03
40	9	100.0	933 21	TERPRAN	Tetrahymena pyriformis	8.17e+03
41	9	100.0	963 23	BTGCEB3	Bos taurus membrane gu	8.17e+03
42	9	100.0	1001 18	AF047171	Providencia stuartii p	8.17e+03
43	9	100.0	1011 17	SMTEMA0GE	S. marcescens beta-lact	8.17e+03
44	9	100.0	1047 22	AF031352	Euplotes aediculatus E	8.17e+03
45	9	100.0	1050 18	AF051814	Escherichia coli strai	8.17e+03
46	9	100.0	1071 17	KPY17583	Klebsiella pneumoniae	8.17e+03
47	9	100.0	1079 17	SUN131631	Streptococcus uberis.	8.17e+03
48	9	100.0	1100 25	I31987	Sequence 2 from Patent	8.17e+03
49	9	100.0	1106 25	I17239	Sequence 4 from patent	8.17e+03
50	9	100.0	1110 11	E08994	DNA encoding an immuno	8.17e+03
51	9	100.0	1134 21	CEI236896	Caenorhabditis elegans	8.17e+03
52	9	100.0	1140 23	AF053028	Panthera tigris altaic	8.17e+03
53	9	100.0	1204 23	SSPAG201	Sus scrofa pregnancy-a	8.17e+03
54	9	100.0	1212 25	A04190	Artificial pregnancy fo	8.17e+03
55	9	100.0	1234 17	RAU59717	Rickettsia akari citra	8.17e+03
56	9	100.0	1241 21	TCR6226	Trypanosoma cruzi mpx	8.17e+03
57	9	100.0	1253 17	PFPCPA	P. furiosus cpkA gene.	8.17e+03
58	9	100.0	1297 23	CFPCD4	C. familiaris mRNA for	8.17e+03
59	9	100.0	1301 23	BTY15995	Bos taurus PIR-1 gene.	8.17e+03
60	9	100.0	1305 23	BTCONGL	B. taurus mRNA for cong	8.17e+03
61	9	100.0	1310 25	I09391	Sequence 24 from Paten	8.17e+03
62	9	100.0	1346 27	AV577768	A. velutinum chloroplas	8.17e+03
63	9	100.0	1346 21	MSE223384	Manduca sexta mRNA for	8.17e+03
64	9	100.0	1360 21	CELCEREPP	C. elegans Cerep3 repa	8.17e+03
65	9	100.0	1365 18	AF106032	Bacillus cerearothermop	8.17e+03
66	9	100.0	1368 22	CEU10327	Caenorhabditis elegans	8.17e+03
67	9	100.0	1400 25	DREHOD4	Danio rerio mRNA for H	8.17e+03
68	9	100.0	1412 25	A32415	H. sapiens PAI-2 CDNA 1	8.17e+03
69	9	100.0	1429 27	AGITITS76	A. githago L. nrDNA (ox	8.17e+03
70	9	100.0	1457 17	ECOCOLEZIC	Escherichia coli plasm	8.17e+03
71	9	100.0	1477 17	EIN9981	Enterobacter intermei	8.17e+03
72	9	100.0	1482 25	I14026	Sequence 5 from patent	8.17e+03
73	9	100.0	1489 18	AF114501	Uncultured gamma prote	8.17e+03
74	9	100.0	1544 25	E13732	CDNA encoding rat CC c	8.17e+03
75	9	100.0	1555 17	CHRDAA	C. trachomatis 16S rbo	8.17e+03
76	9	100.0	1561 24	AB012882	Cyprinus carpio mRNA f	8.17e+03
77	9	100.0	1563 24	AF048817	Xenopus laevis clone 4	8.17e+03
78	9	100.0	1569 24	CRANEUR	Carrasius auratus neur	8.17e+03
79	9	100.0	1573 23	BOVINOPHOS	Bovine inositol polyph	8.17e+03
80	9	100.0	1578 23	OC09355	Oryctolagus cuniculus	8.17e+03
81	9	100.0	1614 25	ARO22629	Sequence 4 from patent	8.17e+03
82	9	100.0	1646 25	A32139	H. sapiens cyclin A gen	8.17e+03
83	9	100.0	1652 25	I01971	Sequence 2 from Patent	8.17e+03
84	9	100.0	1558 25	E13663	CDNA encoding cytocho	8.17e+03
85	9	100.0	1671 23	RABUGTZBB	Oryctolagus cuniculus	8.17e+03
86	9	100.0	1689 21	PTALPHPT1	P. tetraurelia alphaPT1	8.17e+03

C	87	9	100.0	1698 21	LMNACHRB	Locusta migratoria mRN	8.17e+03
C	88	9	100.0	1700 21	APLAP5H	Aplysia californica AP	8.17e+03
C	89	9	100.0	1704 25	A00769	A.hypsoaea gene for st	8.17e+03
C	90	9	100.0	1716 21	LPACT11	L.polyphemus mRNA for	8.17e+03
C	91	9	100.0	1758 25	E02726	DNA encoding human ser	8.17e+03
C	92	9	100.0	1758 25	E02715	CDNA sequence coding f	8.17e+03
C	93	9	100.0	1761 11	E10761	Variant cDNA encoding	8.17e+03
C	94	9	100.0	1786 23	BOVCYPB	Bovine cytochrome P-45	8.17e+03
C	95	9	100.0	1787 25	I50944	Sequence 13 from paten	8.17e+03
C	96	9	100.0	1790 21	CEICOL2G	C.elegans (nematode) c	8.17e+03
C	97	9	100.0	1791 25	E01382	CDNA encoding mature t	8.17e+03
C	98	9	100.0	1797 25	I00062	Sequence 3 from patent	8.17e+03
C	99	9	100.0	1814 24	GGU30520	Gallus gallus P311 POU	8.17e+03
C	100	9	100.0	1830 25	A46379	Sequence 1 from patent	8.17e+03
C	101	9	100.0	1863 25	I17765	Sequence 2 from patent	8.17e+03
C	102	9	100.0	1863 25	AF026812	Butyrylcholinesterase	8.17e+03
C	103	9	100.0	1876 21	SFRACT15B	Strongylocentrotus fra	8.17e+03
C	104	9	100.0	1877 17	ABU09869	Azospirillum brasilens	8.17e+03
C	105	9	100.0	1932 23	AF053483	Felis catus butyrylcho	8.17e+03
C	106	9	100.0	1972 24	CCPMTX13	Quail gene for alpha-t	8.17e+03
C	107	9	100.0	1974 21	NTU80060	Naegleria italica 18S	8.17e+03
C	108	9	100.0	1979 18	AF084366	Lactobacillus bulgari	8.17e+03
C	109	9	100.0	1980 25	I50943	Sequence 12 from paten	8.17e+03
C	110	9	100.0	1980 21	NTU80061	Naegleria jamaicensi 18	8.17e+03
C	111	9	100.0	1984 21	NAU80057	Naegleria andersoni 18	8.17e+03
C	112	9	100.0	1992 17	ECPERABCD	E.coli pery, per, per	8.17e+03
C	113	9	100.0	2006 24	DRU31079	Danio rerio 47 kDa hea	8.17e+03
C	114	9	100.0	2015 25	I04790	Sequence 3 from patent	8.17e+03
C	115	9	100.0	2037 25	A40231	Sequence 1 from patent	8.17e+03
C	116	9	100.0	2046 24	AF031238	Gallus gallus photorec	8.17e+03
C	117	9	100.0	2053 23	MP003906	Macrosclides probosci	8.17e+03
C	118	9	100.0	2053 11	E11452	CDNA encoding plant pr	8.17e+03
C	119	9	100.0	2083 22	AF11868	Entodinium caudatum he	8.17e+03
C	120	9	100.0	2090 21	CEU20902	Caenorhabditis elegans	8.17e+03
C	121	9	100.0	2112 21	OCU09183	Oryctolagus cuniculus	8.17e+03
C	122	9	100.0	2113 21	CEU09415	Caenorhabditis elegans	8.17e+03
C	123	9	100.0	2123 25	GDLAMBR	Chicken mRNA for lamn	8.17e+03
C	124	9	100.0	2136 25	A03759	H.sapiens (pX153) gene	8.17e+03
C	125	9	100.0	2147 21	S73803	HLC-32-32 kDa hyaline	8.17e+03
C	126	9	100.0	2157 21	ECU6825	Encephalozoon cunicu	8.17e+03
C	127	9	100.0	2160 24	CHKCP4502	Chicken phenobarbital-	8.17e+03
C	128	9	100.0	2176 17	SSAGAPD	S.solitaricus acAPD 9	8.17e+03
C	129	9	100.0	2202 24	S76880	XidX-BHL function dom	8.17e+03
C	130	9	100.0	2220 25	A30921	VCAM variant DNA from	8.17e+03
C	131	9	100.0	2223 21	CEGLY5B	Caenorhabditis elegans	8.17e+03
C	132	9	100.0	2225 25	I78692	Sequence 3 from patent	8.17e+03
C	133	9	100.0	2235 25	I09216	Sequence 1 from patent	8.17e+03
C	134	9	100.0	2260 17	R1RTSW56A	Rickettsia tsutsugamus	8.17e+03
C	135	9	100.0	2308 23	TGLHMB	Tachygllossus aculeatus	8.17e+03
C	136	9	100.0	2347 23	BRASPAMTA	B.taurus mitochondrial	8.17e+03
C	137	9	100.0	2350 21	SUSCITLIIA	Strongylocentrotus pur	8.17e+03
C	138	9	100.0	2375 17	ECRFPYRE	E. coli DNA with 2 uni	8.17e+03
C	139	9	100.0	2400 23	PICAPOLP5	Sus scrofa apolipoprot	8.17e+03
C	140	9	100.0	2409 25	A16034	molecule with the acti	8.17e+03
C	141	9	100.0	2425 24	AF082664	Gallus gallus interter	8.17e+03
C	142	9	100.0	2440 11	E11595	Human CDNA encoding en	8.17e+03
C	143	9	100.0	2466 24	AF077838	Xenopus laevis lamln b	8.17e+03
C	144	9	100.0	2494 23	SSU08351	Sus scrofa vascular ce	8.17e+03
C	145	9	100.0	2500 25	I08629	Sequence 30 from Paten	8.17e+03
C	146	9	100.0	2580 25	A20130	Hybrid Insulin/TGF-1 r	8.17e+03
C	147	9	100.0	2599 17	LACOPEP	Lactococcus lactis pep	8.17e+03
C	148	9	100.0	2635 13	RABPRLR	Rabbit prolactin recep	8.17e+03
C	149	9	100.0	2681 25	E12260	CDNA encoding a human	8.17e+03
C	150	9	100.0	2700 24	S72373	Tlx-tailless homolog f	8.17e+03
C	151	9	100.0	2704 23	AF016539	Bos taurus enteric bet	8.17e+03
C	152	9	100.0	2713 23	I13760	Sequence 1 from patent	8.17e+03
C	153	9	100.0	2713 24	CU007694	Columba livia prolacti	8.17e+03
C	154	9	100.0	2723 24	CHREGRA	Gallus gallus EGF/TGF-	8.17e+03
C	155	9	100.0	2742 23	BOVPP36G	Bos taurus (clone Sal3	8.17e+03
C	156	9	100.0	2760 24	AB004650	Chicken mRNA for CENP-	8.17e+03
C	157	9	100.0	2791 23	BOVBARK	Cow beta-adrenergic re	8.17e+03
C	158	9	100.0	2810 23	AB000172	Porcine mRNA for endop	8.17e+03
C	159	9	100.0	2863 24	DB8363	Gallus gallus mRNA for	8.17e+03

C	160	9	100.0	2873 21	DMYEEL5	Drosophila melanogaste	8.17e+03
C	161	9	100.0	2899 25	A44216	Sequence 2 from patent	8.17e+03
C	162	9	100.0	2927 11	E11368	Nucleotide sequence of	8.17e+03
C	163	9	100.0	2939 21	LV033888	Lytechinus variegatus	8.17e+03
C	164	9	100.0	2948 23	OCU21155	Oryctolagus cuniculus	8.17e+03
C	165	9	100.0	2952 24	AB004649	Gallus gallus mRNA for	8.17e+03
C	166	9	100.0	2971 17	MTDNAFLPF	M.thermophilum	8.17e+03
C	167	9	100.0	2989 25	I50971	Sequence 1 from patent	8.17e+03
C	168	9	100.0	3005 21	CCSP70PR	C.capitata gene encodi	8.17e+03
C	169	9	100.0	3020 23	OCU21156	Oryctolagus cuniculus	8.17e+03
C	170	9	100.0	3024 17	PRCCDB	Plasmod KIL18 (from E.	8.17e+03
C	171	9	100.0	3030 24	AF090111	Xenopus laevis (phos E.	8.17e+03
C	172	9	100.0	3104 25	A13388	pPPT18 DNA sequence.	8.17e+03
C	173	9	100.0	3106 23	BRVATPA	B.taurus mRNA for bov1	8.17e+03
C	174	9	100.0	3128 17	ECAL1614	Escherichia coli plasm	8.17e+03
C	175	9	100.0	3144 24	GGCHTRA	G.gallus mRNA for Chit	8.17e+03
C	176	9	100.0	3150 21	OVGS12G	O.vulvulus GST2 gene f	8.17e+03
C	177	9	100.0	3170 18	SPU76218	Streptococcus pneumoni	8.17e+03
C	178	9	100.0	3225 25	I82816	Sequence 13 from paten	8.17e+03
C	179	9	100.0	3239 24	XLGATAG	X.laevis mRNA for GATA	8.17e+03
C	180	9	100.0	3317 21	EMOCHOC1	Silkworm (B.mori) chor	8.17e+03
C	181	9	100.0	3323 17	VIBAGAB	Vibrio sp. agab gene f	8.17e+03
C	182	9	100.0	3348 25	I44532	Sequence 34 from paten	8.17e+03
C	183	9	100.0	3493 24	D49637	Xenopus laevis mRNA fo	8.17e+03
C	184	9	100.0	3517 24	AF012465	Plutoneutes americana	8.17e+03
C	185	9	100.0	3528 25	E13668	CDNA encoding DNA-bind	8.17e+03
C	186	9	100.0	3553 25	I08049	Sequence 5 from patent	8.17e+03
C	187	9	100.0	3643 24	CHKNKATPP	Gallus gallus Na+/K+-A	8.17e+03
C	188	9	100.0	3655 25	I66488	Sequence 8 from patent	8.17e+03
C	189	9	100.0	3681 17	HGRMSG	H.giganteus hgidR & h	8.17e+03
C	190	9	100.0	3699 23	AR021440	Sequence 120 from pate	8.17e+03
C	191	9	100.0	3710 21	DROUMPS	Drosophila melanogaste	8.17e+03
C	192	9	100.0	3750 24	XLBCADH	X.laevis XB-cad mRNA f	8.17e+03
C	193	9	100.0	3754 25	A42478	Sequence 17 from patent	8.17e+03
C	194	9	100.0	3754 25	A37075	Sequence 16 from patent	8.17e+03
C	195	9	100.0	3768 25	A37075	Sequence 18 from patent	8.17e+03
C	196	9	100.0	3775 22	SMHSF3	Schistosoma mansoni he	8.17e+03
C	197	9	100.0	3786 17	TNLAMABOL	T.neopolitana lama gen	8.17e+03
C	198	9	100.0	3820 17	STRATPEFHA	Streptococcus faecalis	8.17e+03
C	199	9	100.0	3821 25	E07844	Genomic DNA of human n	8.17e+03
C	200	9	100.0	3886 27	ATPHXD	A.thaliana phyd gene.	8.17e+03

ALIGNMENTS

RESULT 1
LOCUS 146684
DEFINITION Sequence 663 from patent US 5639612.
ACCESSION 146684
NID 92470649
VERSION 146684.1 GI:2470649
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 26)
AUTHORS Mitsuhashi, M. and Cooper, A.
TITLE Method for detecting polynucleotides with immobilized
poly-nucleotide probes identified based on T.sub.m
JOURNAL Patent: US 5639612-A 663 17-JUN-1997;
FEATURES
Location/Qualifiers
1..26
BASE COUNT 10 a /organism="unknown"
ORIGIN 4 c 6 g 6 t

Query Match 100.0%; Score 9; DB 25; Length 26;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTCAGAA 13
Cp 33 ttctcagaa 25

RESULT 2 146688 26 bp DNA PAT 29-JUN-1997
 LOCUS DEFINITION Sequence 667 from patent US 5639612.
 ACCESION 146688
 NID 92470653
 VERSION 146688.1 GI:2470653
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 26)
 AUTHORS Mitsuhashi, M. and Cooper, A.
 TITLE Method for detecting polynucleotides with immobilized
 polynucleotide probes identified based on T.sud.m
 JOURNAL Patent: US 5639612-A 667 17-JUN-1997;
 FEATURES Location/Qualifiers
 source 1..26
 BASE COUNT 10 a 4 c 6 g 6 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 26;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTCAGAA 13
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 3 132126 132 bp DNA PAT 07-JAN-1997
 LOCUS DEFINITION Sequence 12 from patent US 5585248.
 ACCESION 132126
 NID 91822917
 VERSION 132126.1 GI:1822917
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 132)
 AUTHORS Ashida, M., Kawabata, T., Hirayasu, K. and Tsuchiya, M.
 TITLE Method for assaying activity of prophenoloxidase activating enzyme
 and application thereof
 JOURNAL Patent: US 5585248-A 12 17-DEC-1996;
 FEATURES Location/Qualifiers
 source 1..132
 BASE COUNT 46 a 34 c 21 g 31 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 132;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 TTCTGAGAA 89
 |||||||||
 Oy 25 ttctcagaa 33

RESULT 4 184379 192 bp DNA PAT 20-MAR-1998
 LOCUS DEFINITION Sequence 37 from patent US 5695932.
 ACCESION 184379
 NID 93021899
 VERSION 184379.1 GI:3021899
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 192)
 AUTHORS West, M.D., Shay, J., Wright, M., Blackburn, E.H. and McEachern, M.J.
 TITLE Telomerase activity assays for diagnosing pathogenic infections
 JOURNAL Patent: US 5695932-A 37 09-DEC-1997;
 FEATURES Location/Qualifiers
 source 1..192
 BASE COUNT 58 a 49 c 33 g 52 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 192;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 41 TTCTCAGAA 49
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 5 A59593 198 bp DNA PAT 06-MAR-1998
 LOCUS DEFINITION Sequence 6 from Patent WO9704112.
 ACCESION A59593
 NID 93714903
 VERSION A59593.1 GI:3714903
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 198)
 AUTHORS Poulsen, P.
 TITLE INHIBITION OF GENE EXPRESSION
 JOURNAL Patent: WO 9704112-A 6 06-FEB-1997;
 FEATURES Location/Qualifiers
 source 1..198
 BASE COUNT 64 a 32 c 29 g 73 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 198;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 TTCTCAGAA 40
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 6 BOVC03 242 bp DNA MAM 27-JUN-1997
 LOCUS DEFINITION Bovine gene for conglutinin, exon III.
 ACCESION D25296
 NID 9483418
 VERSION D25296.1 GI:483418
 KEYWORDS conglutinin.
 SEGMENT 3 of 9
 SOURCE Bos taurus mature liver hepatocyte DNA.
 ORGANISM Bos taurus
 REFERENCE 1 (bases 1 to 242)
 AUTHORS Kawasaki, N.
 TITLE Direct Submision
 JOURNAL Submitted (18-NOV-1993) to the DDBJ/EMBL/GenBank databases. Nobuko
 Kawasaki, College of Medical Technology, Kyoto University;
 Kawarumachi 53, Shogoin, Sakyo-ku, Kyoto, Kyoto 606-01, Japan
 (Tel: 075-751-3934, Fax: 075-751-3909)
 REFERENCE 2 (bases 1 to 242)
 AUTHORS Kawasaki, N., Itoh, N. and Kawasaki, T.
 TITLE Gene organization and 5'-flanking region sequence of conglutinin: a
 C-type mammalian lectin containing a collagen-like domain
 JOURNAL Biochem. Biophys. Res. Commun. 198 (2), 597-604 (1994)

MEDLINE 94128104 Location/Qualifiers
 FEATURES
 Source
 1. .242
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /cell_type="hepatocyte"
 /dev_stage="mature"
 /tissue_type="liver"
 join(D25294.1:571. .738,D25295.1:21. .92,21. .23)
 5'UTR
 Intron
 <1. .20
 /number=2
 21. .222
 /number=3
 /product="conglutinin"
 21. .23
 /product="conglutinin"
 24. .83
 /product="conglutinin signal peptide"
 223. .>242
 /note="size: >2.1 kbp"
 /number=3
 BASE COUNT 50 a 67 c 67 g 58 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 23; Length 242;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 101 TTCTCAGAA 109
 |||||
 33 ttctcagaa 25

RESULT 7
 LOCUS BRU65795 246 bp DNA BCT 02-JAN-1997
 DEFINITION Borrelia burgdorferi strain All outer surface protein A (ospA)
 ACCESSION U65795
 NID 91762792
 VERSION U65795.1 GI:1762792
 KEYWORDS
 SOURCE Lyme disease spirochete.
 ORGANISM Borrelia burgdorferi
 Eubacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 REFERENCE 1 (bases 1 to 246)
 Mathiesen,D.A., Oliver,Jr., J.H., Kolbert,C.P., Tullison,E.D.,
 Johnson,B.J.B., Campbell,G.L., Mitchell,P.D., Reed,R.D.,
 Telford,I.I.S.R., Anderson,J.F., Lane,R.S. and Persing,D.H.
 Genetic Diversity of Borrelia burgdorferi
 Unpublished
 JOURNAL 2 (bases 1 to 246)
 REFERENCE Mathiesen,D.A. and Persing,D.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-AUG-1996) Experimental Path., Mayo Clinic, Rochester,
 MN 55905, USA

FEATURES
 Source
 Location/Qualifiers
 1. .246
 /organism="Borrelia burgdorferi"
 /strain="All"
 /db_xref="taxon:139"
 1. .246
 /gene="ospA"
 <1. .>246
 /gene="ospA"
 /codon_start=2
 /transl_table=1
 /product="outer surface protein A"
 /protein_id="AA839370.1"
 /db_xref="PI:91762793"
 /db_xref="GI:1762793"
 /translation="ACKQAVSSLDKKNYSVDLPGEKVLVSEKDKDKGYSIMATV
 DNLEKTSKKNNGSVLSEGVKADSKVKLTVEEDLST"
 BASE COUNT 100 a 33 c 55 g 56 t

Query Match 100.0%; Score 9; DB 17; Length 246;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 157 TTCTGAGAA 165
 |||||
 25 ttctgagaa 33

RESULT 8
 LOCUS CYRS 250 bp DNA VRT 05-NOV-1991
 DEFINITION Carp (C.carpio) repetitive sequence, clone pCchr-3.
 ACCESSION M19418
 NID 9213062
 VERSION M19418.1 GI:213062
 KEYWORDS repeat region.
 SOURCE Carp (C.carpio, strain Bangkok, adult) erythrocyte DNA, clone pCchr-3.
 ORGANISM Cyprinus carpio
 Eukaryota; Mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae;
 Cyprinus.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Datta,U., Dutta,P. and Mandal,R.K.
 TITLE Cloning and characterization of a highly repetitive fish nucleotide
 sequence
 JOURNAL Gene 62: 331-336 (1988)
 COMMENT 88212186
 DRAFT entry and printed copy of sequence for [1] kindly provided by
 R.K.Mandal, 02-JUN-1988.
 LOCATION/Qualifiers
 1. .250
 /organism="Cyprinus carpio"
 /db_xref="taxon:7962"
 repeat_region
 1. .250
 /note="repetitive sequence"
 BASE COUNT 70 a 47 c 44 g 89 t
 ORIGIN
 1 bp upstream of HindIII site.

Query Match 100.0%; Score 9; DB 24; Length 250;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 TTCTCAGAA 127
 |||||
 33 ttctcagaa 25

RESULT 9
 LOCUS APA011313 256 bp DNA VRT 21-SEP-1998
 DEFINITION Acrossochellus paradoxus HindIII satellite DNA, isolate Wa23.
 ACCESSION AJ011313
 NID 93641827
 VERSION AJ011313.1 GI:3641827
 KEYWORDS satellite.
 SOURCE Acrossochellus paradoxus.
 ORGANISM Acrossochellus paradoxus.
 Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 Cyprinidae; Cyprinidae; Acrossochellus.
 REFERENCE 1 (bases 1 to 256)
 AUTHORS Chiang,T.Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1998) Chiang T.Y., Department of Biology,
 Cheng-kung University, Tainan, university Rd., TAIWAN
 2 (bases 1 to 256)
 REFERENCE Chiang,T.Y.
 TITLE Cloning HindIII satellite DNA of Acrossochellus paradoxus
 JOURNAL Unpublished
 FEATURES
 Location/Qualifiers
 1. .256
 source

/organism="Acrossochellus paradoxus"
/isolate="Ma23"
/db_xref="taxon:76593"
1. .256
/note="hindiii satellite DNA"
/citation={2}

BASE COUNT 77 a 51 c 38 g 90 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 256;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 TTCACGAA 71
|||||
Cp 33 tttcagaa 25

RESULT 10
ID E01842 standard: RNA; VRL: 303 BP.
AC E01842.1
SV E01842.1
NI d1098679
DT 07-OCT-1997 (Rel. 52, Created)
DT 07-OCT-1997 (Rel. 52, Last updated, Version 1)
DE Genomic RNA of hop stunt viroid type HSV-C.
KW JP 1989040000-A/1.
OC Hop stunt viroid
OS Viroids.
RN [1]
RP 1-303
RA Yomo E., Sano T., Kudo H., Sugimoto N.;
RL "METHOD FOR FRACTIONATING AND DETECTING HOP STUNT VIROIDS";
RL Patent number JP 1989040000-A/1, 10-FEB-1989.
RL UKI GOSHI KOGIO CO LTD.
RL HOP STUNT VIROID
CC OS Hop stunt viroid
CC PN JP 1989040000-A/1
CC PD 10-FEB-1989
CC PF 05-AUG-1987 JP 1987194377
CC PI YOMO EISHIRO, SANO TERUO, KUDO HIROSHI, SUGIMOTO NOBUTAKA
CC PC C12Q1/70,C12N15/00,C12Q1/68,G01N33/50;
CC CC strandedness: Single;
CC CC topology: Circular;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC *source: strain=HSV-C;
FH Key Location/Qualifiers
FT source 1. .303
FT /organism="Hop stunt viroid"
FT /db_xref="taxon:12893"
FT Sequence 303 BP: 64 A; 87 C; 81 G; 71 T; 0 other;
SQ

Query Match 100.0%; Score 9; DB 11; Length 303;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 92 TTCACGAA 100
|||||
Cp 33 tttcagaa 25

RESULT 11
LOCUS CNMTB 309 bp DNA MAM 16-DEC-1996
DEFINITION Conepatus leuconotus mitochondrial DNA.
ACCESSION L27299
NID 9439414
VERSION L27299.1 GI:439414
KEYWORDS
SOURCE Conepatus leuconotus
ORGANISM Mitochondrion Conepatus leuconotus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Carnivora; Fissipedia; Mustelidae; Conepatus.

REFERENCE 1 (bases 1 to 309)
AUTHORS Drago,J.W., Bradley,R.D., Honeycutt,R.L. and Templeton,J.W.
TITLE Phylogenetic Relationships Among the Skunks: A Molecular
JOURNAL J. Mamm. Evol. 1, 255-267 (1993)
FEATURES
source 1. .309
/organism="Conepatus leuconotus"
/mitochondrion
/db_xref="taxon:30543"
BASE COUNT 90 a 82 c 45 g 92 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 232 TTCACGAA 240
|||||
Cp 33 tttcagaa 25

RESULT 12
LOCUS REECYTB1 331 bp DNA MAM 15-JUL-1997
DEFINITION Rhinolophus ferrumequinum 5' cytochrome b (cytb) gene,
mitochondrial gene encoding mitochondrial protein, partial cds.
ACCESSION U95513
NID 92253673
VERSION U95513.1 GI:2253673
KEYWORDS
SEGMENT 1 of 2
SOURCE Rhinolophus ferrumequinum.
ORGANISM Mitochondrion Rhinolophus ferrumequinum
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Chiroptera; Microchiroptera;
Rhinolophidae; Rhinolophus.
REFERENCE 1 (bases 1 to 331)
AUTHORS Barrett,E.M., Deaville,R., Burland,T.M., Bruford,M.W., Jones,G.,
Racey,P.A. and Wayne,R.K.
TITLE DNA answers the call of pipistrelle bat species
JOURNAL Nature 387 (6629), 138-139 (1997)
MEDLINE 97289376
REFERENCE 2 (bases 1 to 331)
AUTHORS Barrett,E.M., Deaville,R.C., Burland,T.M., Bruford,M.W., Jones,G.,
Racey,P.A. and Wayne,R.K.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-1997) Conservation Genetics, Institute of
Zoology, Regents Park, London NW1 4RY, England
FEATURES
source 1. .331
/organism="Rhinolophus ferrumequinum"
/mitochondrion
/db_xref="taxon:59479"
CDS
1. .331
/gene="cytb"
/codon_start=1
/transl_table=2
/product="cytochrome b"
/protein_id="AAC48750.1"
/db_xref="PID:92253676"
/db_xref="GI:2253676"
/translation="PLEKIIINDSFVDPAPSSISSWMNFGSLGICLAIOILTGLFLA
MHTSDPATAFHSVTHICRDVNYGWVRLYHANGASMFICLFHVGRCIYGSYTES
ETNNIGIT"

BASE COUNT 88 a 110 c 49 g 84 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 331;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 301 TTCACGAA 309
|||||

Cp 33 ttctcagaa 25

```

RESULT 13
LOCUS RAU33923 367 bp DNA BCT 13-MAR-1997
DEFINITION Rickettsia australis citrate synthase gene, partial cds.
ACCESSION U33923
NID 91015949
VERSION U33923.1 GI:1015949
KEYWORDS
SOURCE
ORGANISM Rickettsia australis strain-JC.
            Rickettsia australis
            Eubacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Rickettsiella; Rickettsia.
REFERENCE 1 (bases 1 to 367)
AUTHORS Higgins,J.A., Radulovic,S., Schrieler,M.E. and Azad,A.F.
TITLE Rickettsia felis: a new species of pathogenic rickettsia isolated
        from cat fleas
JOURNAL J. Clin. Microbiol. 34 (3), 671-674 (1996)
MEDLINE 97060114
REFERENCE 2 (bases 1 to 367)
AUTHORS Higgins,J.A., Radulovic,S. and Schrieler,M.E.
TITLE Direct Submission
JOURNAL Submitted (15-AUG-1995) Abdu Azad, Microbiology, University of
        Maryland at Baltimore, School of Medicine, 655 W Baltimore,
        Baltimore, MD 21201, USA
FEATURES
    source
        1..367
        /organism="Rickettsia australis"
        /strain="JC"
        /db_xref="taxon:787"
        <1..>367
        /codon_start=2
        /product="citrate synthase"
        /protein_id="AA049576.1"
        /db_xref="PID:91015950"
        /db_xref="GI:1015950"
        /translation="AHGAGNEAVINMLKEIGSENPPIKAYKDKNDPRLMGFGR
        VKNYNDPRAALKECKKAVLKEGLQENPLDIAIEALAKDEYFERKLYPRVD
        SHSGIYKAMGIRISOMSTVL"
BASE COUNT 129 a 62 c 74 g 102 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 367;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 TTCTGAGAA 63
|||||
OY 25 ttctgagaa 33

RESULT 14
LOCUS OA047049 374 bp DNA MAM 12-DEC-1996
DEFINITION Ovis aries TG dinucleotide repeat region.
ACCESSION U47049
NID 91199625
VERSION U47049.1 GI:1199625
KEYWORDS
SOURCE sheep.
ORGANISM Ovis aries
            Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 374)
AUTHORS Hawken,R.J. and Maddox,J.F.
TITLE An ovine dinucleotide repeat polymorphism at the RH1 locus
JOURNAL Anim. Genet. 27 (3), 219-220 (1996)
MEDLINE 96314793
REFERENCE 2 (bases 1 to 374)
AUTHORS Hawken,R.J.
TITLE Direct Submission

```

```

JOURNAL Submitted (24-JAN-1996) Rachel J. Hawken, Centre for Animal
        Biotechnology, The University of Melbourne, Parkville, Victoria
        3052, Australia
FEATURES
    source
        1..374
        /organism="Ovis aries"
        /db_xref="taxon:9940"
        /chromosome="9"
        repeat_region
            274..299
            /rpt_family="TG"
            /rpt_type="tandem"
BASE COUNT 116 a 53 c 67 g 138 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 74 TTCTCAGAA 82
|||||
Cp 33 ttctcagaa 25

RESULT 15
LOCUS RA041752 383 bp DNA BCT 13-MAR-1997
DEFINITION Rickettsia akari citrate synthase gene, partial cds.
ACCESSION U41752
NID 9145792
VERSION U41752.1 GI:1145792
KEYWORDS
SOURCE Rickettsia akari strain-Kaplan.
ORGANISM Rickettsia akari
            Eubacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Rickettsiella; Rickettsia.
REFERENCE 1 (bases 1 to 383)
AUTHORS Higgins,J.A., Radulovic,S., Schrieler,M.E. and Azad,A.F.
TITLE Rickettsia felis: a new species of pathogenic rickettsia isolated
        from cat fleas
JOURNAL J. Clin. Microbiol. 34 (3), 671-674 (1996)
MEDLINE 97060114
REFERENCE 2 (bases 1 to 383)
AUTHORS Azad,A., Higgins,J.A., Radulovic,S. and Schrieler,M.E.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1995) Abdu Azad, Microbiology and Immunology,
        University of Maryland at Baltimore, 655 W. Baltimore Street,
        Baltimore, MD 21201, USA
FEATURES
    source
        1..383
        /organism="Rickettsia akari"
        /strain="Kaplan"
        /db_xref="taxon:786"
        <1..>383
        /codon_start=1
        /product="citrate synthase"
        /protein_id="AA049577.1"
        /db_xref="PID:9145793"
        /db_xref="GI:1145793"
        /translation="WGAHAGNEAVINMLKEIGSENPPIKAYKDKNDPRLMGF
        GHRVKNYNDPRAASVLEKCKKAVLKEGLQENPLDIAIEALAKDEYFERKLYPR
        NYDRISGIYKAMGIRISOMFTVYTR"
BASE COUNT 134 a 58 c 77 g 114 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 383;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 TTCTGAGAA 71
|||||
OY 25 ttctgagaa 33

```

[illegible]

FEATURES	SOURCE
location/Qualifiers	1. :385
/organism="Entamoeba invadens"	
/strain="rp-1"	
/isolate="INVD1"	
/db_xref="taxon:33085"	
/clone="INVD1"	
/dev_stage="Trophozoite"	
/cell_type="Trophozoite"	
/clone_lib="ribosomal DNA"	
<1. :>385	
rrna	
/product="ribosomal RNA"	
BASE COUNT	176 a 76 g 90 t
ORIGIN	43 c

Query Match	100.0%;	Score 9;	DB 21;	Length 385;
Best Local Similarity	100.0%;	Pred. No. 8.17e+03;		
Matches	9;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
Db	310	TTCTGAGAA	318	
OY	25	ttctgagaa	33	

[illegible]

```

tRNA      <1..>396
           /gene="16S rRNA"

```

```

/note="PCR product"
/product="16S ribosomal RNA"
gene
1. .396
/gene="16S rRNA"
BASE COUNT      102 a      89 c      134 g      69 t      2 others
ORIGIN
Query Match      100.0%; Score 9; DB 17; Length 396;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
Db      253 TTCGAGAA 261
|||||||
Qy      25 ttctgagaa 33

```

RESULT		18				
LOCUS	MSU75307	399 bp	mRNA	INV	26-JUN-1998	
DEFINITION	Manduca sexta cellular retinoic acid binding protein (mCRABP) mRNA,					
	complete cds.					
ACCESSION	U75307					
NID	G3115356					
VERSION	U75307.1 GI:3115356					
KEYWORDS	.					
SOURCE	tobacco hornworm.					
ORGANISM	Manduca sexta					
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;					
	Pterygota; Lepidoptera; Sphingodea; Spingiidae; Manduca.					
REFERENCE	1 (bases 1 to 399)					
AUTHORS	Mansfield,S.G., Cammer,S., Alexander,S.C., Muehleisen,D.P., Gray,R.S., Tropsha,A. and Bollenbacher,W.E. Molecular cloning and characterization of an invertebrate cellular retinolic acid binding protein <i>Proc. Natl. Acad. Sci. U.S.A.</i> 95 (12), 6825-6830 (1998) JOURNAL MEDLINE REFERENCE AUTHORS					
TITLE	JOURNAL MEDLINE REFERENCE AUTHORS					
TITLE	JOURNAL MEDLINE REFERENCE AUTHORS					
JOURNAL	JOURNAL MEDLINE REFERENCE AUTHORS					
Medline	JOURNAL MEDLINE REFERENCE AUTHORS					
Reference	JOURNAL MEDLINE REFERENCE AUTHORS					
Authors	JOURNAL MEDLINE REFERENCE AUTHORS					
Title	JOURNAL MEDLINE REFERENCE AUTHORS					
Journal	JOURNAL MEDLINE REFERENCE AUTHORS					
Submitted	(16-OCT-1996) Biology, UNC Chapel Hill, 333 Wilson Hall, Chapel Hill, NC 27599 USA					

gene	/organism="Manduca sexta"	
	/db_xref="taxon:7130"	
	1. .399	
CDS	/gene="mCRAB"	
	1. .399	
	/gene="mCRAB"	
	/note="lipid binding protein"	
	/codon_start=1	
	/product="cellular retinoic acid binding protein"	
	/protein_id="M24317.1"	
	/db_xref="PID:g3115357"	
	/db_xref="GI:3115357"	
	/translation="MEFVQKRYKMSSENFDEPKATGVLITRKANAAYPTVELRQ	
	EGDGIYLVTSSTFKTTEMKRFRPEEPEDERADGAKVSVCTFEQNTLKYQKADGEE	
	VTYIRREGPEEMKVMNAKDYTCIRYKVO"	
BASE COUNT	104 a	102 c
ORIGIN	121 g	72 t
Query Match	100.0%;	Score 9;
Best Local Similarity	100.0%;	DB 21; Length 399;
Matches	9; Conservative	Pred. No. 8.17e+03;
	0; Mismatches	0; Indels
		Gaps 0;
DB	36 TTCTGAGAA	44
QY	25 ttctcgagaa	33

RESULT	19			
LOCUS	AF044660	402 bp	DNA	MAM
DEFINITION	Rhinolophus hipposideros cytochrome b gene, mitochondrial gene			

```

ACCESSION      encoding mitochondrial protein, partial cds.
AF044660
NID            94159980
VERSION        AF044660.1 GI:4159980
KEYWORDS
SOURCE
ORGANISM
Rhizophus hipposideros.
Mitochondrion Rhizophus hipposideros
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Chiroptera; Microchiroptera; Rhinolophidae; Rhinolophus.
REFERENCE
1 (bases 1 to 402)
Bautista,J.M., Alvarez,Y. and Juste,J.
Mitochondrial phylogeography and morphological recurrence in
African fruitbats
unpublished
JOURNAL
2 (bases 1 to 402)
Bautista,J.M., Alvarez,Y. and Juste,J.
Direct Submission
Submitted (26-JAN-1998) Biología y Biología Molecular IV,
Universidad Complutense de Madrid, Facultad de Veterinaria, Ciudad
Universitaria s/n, Madrid 28040, Spain
FEATURES
Source
Location/Qualifiers
1..402
/organism="Rhinolophus hipposideros"
/mitochondrion
/db_xref="taxon:77218"
1..>402
/codon_start=1
/product="cytochrome b"
/protein_id="AAD05413.1"
/db_xref="PID:94159981"
/db_xref="GI:4159981"
/translation="MTNRKSHPLEKRIINDSFIDLPMSSTSSMWNFGSLIGVCLATO
ILTCFLPMHYTSDTDRAHSHVTHICDQVNYKIRLIRHANGAMFICLFLHNGRI
YGSYTSFETNNIGILFAYMATAPKGYVP"
BASE COUNT    112 a 126 c 57 g 107 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 402;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 325 TTCGAGAA 333
|||||
Cp 33 ttctcagaa 25

RESULT 20
LOCUS      AF114520 436 bp DNA BCT 12-MAR-1999
DEFINITION unclutred gamma proteobacterium 400m-ATT-15
gene, complete sequence.
ACCESSION  AF114520
NID        94406325
VERSION    AF114520.1 GI:4406325
KEYWORDS
SOURCE
unclutred gamma proteobacterium 400m-ATT-15.
unclutred gamma proteobacterium 400m-ATT-15
Bacteria; Proteobacteria; gamma subdivision; environmental samples.
REFERENCE
1 (bases 1 to 436)
Acinas,S.G., Anton,J. and Rodriguez-Valera,F.
Diversity of free-living and attached bacteria in offshore western
Mediterranean waters as depicted by analysis of genes encoding 16S
rRNA
JOURNAL
Appl. Environ. Microbiol. 65 (2), 514-522 (1999)
MEDLINE
99124582
REFERENCE
2 (bases 1 to 436)
Acinas,S.G.
Direct Submission
Submitted (23-NOV-1998) Microbiology, University Miguel Hernandez,
Ctra. Valencia Km 87, Alicante, Alicante 03550, Spain
FEATURES
Source
Location/Qualifiers
1..436
/organism="unclutred gamma proteobacterium 400m-ATT-15"

```

```

/Isolate="400m ATT"
/db_xref="taxon:86492"
/clone="15"
/note="uncultured attached marine bacterium recovered from
depth of 400m"
<1..>436
/note="sequenced with reverse primer"
/product="16S ribosomal RNA"
BASE COUNT    118 a 93 c 137 g 88 t
ORIGIN
Query Match 100.0%; Score 9; DB 18; Length 436;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 278 TTCGAGAA 286
|||||
Cp 25 ttctcagaa 33

RESULT 21
LOCUS      AF063883 443 bp DNA INV 20-OCT-1998
DEFINITION Drosophila teissleri copia retrotransposon, 5' long terminal repeat
and untranslated leader region.
ACCESSION  AF063883
NID        93769468
VERSION    AF063883.1 GI:3769468
KEYWORDS
SOURCE
Drosophila teissleri.
Drosophila teissleri.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 443)
Jordan,I.K. and McDonald,J.F.
Inter-element selection in the regulatory region of the copia
retrotransposon
J. Mol. Evol. (1998) In press
JOURNAL
2 (bases 1 to 443)
Jordan,I.K. and McDonald,J.F.
Evolution of the copia retrotransposon in the Drosophila
melanogaster species subgroup
Mol. Biol. Evol. (1998) In press
JOURNAL
3 (bases 1 to 443)
Jordan,I.K. and McDonald,J.F.
Direct Submission
Submitted (08-MAY-1998) Genetics, University of Georgia, 1057 Green
Street, Athens, GA 30602, USA
FEATURES
Source
Location/Qualifiers
1..443
/organism="Drosophila teissleri"
/db_xref="taxon:7243"
<1..>443
/note="5' long terminal repeat LTR and untranslated leader
region UTR"
repeat_region
<1..>443
/rpt_family="copia retrotransposon"
/rpt_type="dispersed"
BASE COUNT    156 a 65 c 60 g 162 t
ORIGIN
Query Match 100.0%; Score 9; DB 22; Length 443;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 402 TTCGAGAA 410
|||||
Cp 33 ttctcagaa 25

RESULT 22
LOCUS      SHPMICB 470 bp DNA NAM 07-JUL-1994
DEFINITION Ovis aries DNA microsatellite.

```

ACCESSION L34278
MID 9508588
VERSION L34278.1 GI:508588
KEYWORDS PCR primer; microsatellite.
SOURCE Ovis aries DNA.
ORGANISM Ovis aries
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
REFERENCE 1 (bases 1 to 470)
AUTHORS Hulme,D.J., Silk,J.P., Redwin,J. and Beh,K.J.
TITLE Ten polymorphic ovine microsatellites
JOURNAL Unpublished (1994)
FEATURES
source Location/Qualifiers
1..470
/organism="Ovis aries"
/db_xref="taxon:9940"
complement(235..258)
/note="PCR primer to amplify microsatellite; PCR primer
binding site; putative"
satellite 286..350
/note="compound microsatellite; putative"
/rpt_type="other
345..369
/standard_name="reverse PCR primer"
/note="PCR primer binding site; putative"
BASE COUNT 117 a 96 c 142 g 113 t 2 others
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 470;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 213 TTCGAGAA 221
|||||
OY 25 ttctgagaa 33
RESULT 23
LOCUS L0004669 479 bp mRNA 03-JUN-1995
DEFINITION Lymantria dispar Ld40 chorion protein mRNA, partial cds.
ACCESSION U04669
MID 9454066
VERSION U04669.1 GI:454066
KEYWORDS
SOURCE Lymantria dispar.
ORGANISM Lymantria dispar.
Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Lepidoptera; Noctuidae;
Lymantriidae; Lymantria.
REFERENCE 1 (bases 1 to 479)
AUTHORS Leclerc,R.F. and Regier,J.C.
TITLE Evolution of chorion gene families in lepidoptera: characterization
of 15 CDNA's from the gypsy moth
JOURNAL J. Mol. Evol. 39 (3), 244-254 (1994)
MEDLINE 95018300
REFERENCE 2 (bases 1 to 479)
AUTHORS Leclerc,R.F.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1993) Robert F. Leclerc, University of Maryland,
Center for Agricultural Biotechnology, Agriculture/Life Sciences
Surge Bldg., College Park, MD 20742-3351, USA
FEATURES
source Location/Qualifiers
1..479
/organism="Lymantria dispar"
/db_xref="taxon:13123"
/clone="Ld40"
/clone_lib="Choriongenic follicular cDNA"
/sex="female"
/tissue_type="choriongenic follicles"
/dev_stage="pupae"
1..420
/codon_start=1
CDS

/product="chorion protein"
/protein_id="AA67869.1"
/db_xref="pid:9454067"
/db_xref="GI:454067"
/translation="LEAPFAMDAPCCGGLSPADSPSTSGALPVSSASAIAPVGLA
VASENVEGILAAAGLEPFVGTGVEGILPTAGAGAVYHSCGDGINMTSRDAFAFR
FADAVGIGLAGYGLEVPILNAPALGYRAGWKGCCGL"
polyA_site 479
/note="8 A nucleotides"
BASE COUNT 77 a 132 c 129 g 141 t
ORIGIN
Query Match 100.0%; Score 9; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 138 TTCGAGAA 146
|||||
OY 25 ttctgagaa 33
RESULT 24
LOCUS CHRCW8 491 bp DNA VRT 09-NOV-1995
DEFINITION Gallus domesticus DNA microsatellite marker MCW8.
ACCESSION L40071
MID 91063495
VERSION L40071.1 GI:1063495
KEYWORDS CA-microsatellite; polymorphic microsatellite.
SOURCE Gallus gallus (clone: K171) (tissue library: lambda ZAP) female
adult blood DNA.
ORGANISM Gallus gallus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 491)
AUTHORS Groenen,M.A.M., Crooijmans,R.P.M. and van der Poel,J.J.
TITLE Preliminary linkage map of the chicken (Gallus domesticus) genome
based on microsatellite markers: 77 new markers mapped
JOURNAL Unpublished (1996)
FEATURES
source Location/Qualifiers
1..491
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="K171"
/dev_stage="adult"
/germline
/sex="female"
/tissue_type="blood"
/tissue_lib="lambda ZAP"
primer_bind 221..244
repeat_region 315..342
/rpt_family="CA-microsatellite"
primer_bind 419..443
BASE COUNT 90 a 72 c 123 g 206 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 146 TTCGAGAA 154
|||||
OY 25 ttctgagaa 33
RESULT 25
LOCUS BOVCYT84 496 bp DNA MAM 31-AUG-1998
DEFINITION Bovine cytochrome b-5 gene, exon 6 and complete cds.
ACCESSION L22966
MID 9387581
VERSION L22966.1 GI:387581
KEYWORDS cytochrome b5; membrane-bound protein.
SEGMENT 4 of 4

SOURCE Bos taurus.
ORGANISM Bos taurus.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 496)
AUTHORS Cristiano,R.J., Giordano,S.J. and Steggles,A.W.
TITLE The isolation and characterization of the bovine cytochrome b5 gene, and a transcribed pseudogene

JOURNAL Genomics 17 (2), 348-354 (1993)
MEDLINE 94010928
FEATURES
source Location/Qualifiers
1..496
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="liver"
order(M63328.1:4162..5786.1..44)
/number=5
Intron
/number=5
exon
45..496
/number=6
/product="cytochrome b-5"
BASE COUNT 134 a 127 c 86 g 149 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 132 TTCTGAGAA 140
|||||
CP 33 ttctcagaa 25

RESULT 26
LOCUS HEF180 536 bp mRNA VRT 26-OCT-1995
DEFINITION Heterodontus francisci tcr beta. gene.
ACCESSION U47454
NID 91032343
VERSION L47454.1 GI:1032343
KEYWORDS T cell receptor beta chain.
SOURCE Heterodontus francisci (tissue library: lambda zap) adult spleen
CDNA to mRNA.
ORGANISM Heterodontus francisci
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Chondrichthyes; Elasmobranchii; Heterodontiformes; Heterodontidae; Heterodontus.
REFERENCE 1 (bases 1 to 536)
AUTHORS Hawke,N.A., Rast,J.P. and Litman,G.W.
TITLE Extensive Diversity of Transcribed TCR Beta in a Phylogenetically Primitive Vertebrate
JOURNAL Unpublished (1993)
FEATURES
source Location/Qualifiers
1..536
/organism="Heterodontus francisci"
/db_xref="taxon:7792"
/dev_stage="adult"
/tissue_type="spleen"
/tissue_lib="lambda zap"
115..536
mat_peptide /partial
BASE COUNT 122 a 149 c 150 g 115 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 536;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 TTCTGAGAA 40
|||||
QY 25 ttctcagaa 33

RESULT 27
LOCUS BS16SRMU 578 bp DNA BCT 04-SEP-1991

DEFINITION B.subtilis (type26 mut) 16S rRNA gene, 5'end.
ACCESSION X55333
NID 939778
VERSION X55333.1 GI:39778
KEYWORDS 16S ribosomal RNA.
SOURCE 16S ribosomal RNA.
ORGANISM Bacillus subtilis.
Bacillus subtilis.
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 578)
AUTHORS Jarvis,E.D., Cheng,S. and Rudner,R.
TITLE Genetic structure and DNA sequences at junctions involved in the rearrangements of Bacillus subtilis strains carrying the trpE26 mutation

JOURNAL Genetics 126 (4), 785-797 (1990)
MEDLINE 91169259
FEATURES
source Location/Qualifiers
1..578
/organism="Bacillus subtilis"
/strain="trpE26-type"
/db_xref="taxon:1423"
/clone="pJr421"
1..221
/note="ORF"
CDS
/codon_start=3
/transl_table=11
/protein_id="CAA39032.1"
/db_xref="PID:9388215"
/db_xref="GI:388215"
/db_xref="SPTREMBL:004384"
/translation="FRFRFRISMTDVELYPFMANSKAVSIISFSTFNGSTLSEP
GGILGKGYORLPDRNERTGLSPRIIF"
230..235
/note="p1 promoter"
253..258
/note="p1 promoter"
327..332
/note="p2 promoter"
350..355
/note="p2 promoter"
537..578
/note="rind ribosomal operon"
/product="16S rRNA"
BASE COUNT 175 a 113 c 126 g 164 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 578;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 278 TTCTGAGAA 286
|||||
QY 25 ttctcagaa 33

RESULT 28
LOCUS I07689 583 bp PAT 14-NOV-1994
DEFINITION Sequence 12 from Patent EP 0364255.
ACCESSION I07689
NID 9589714
VERSION I07689.1 GI:589714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 583)
AUTHORS Caskey,C.T., Chamberlain,J.S., Glibbs,R.A., Rainer,J.E. and Nguyen,P.N.
TITLE Multiplex genomic DNA amplification for deletion detection
JOURNAL Patent: EP 0364255-A2 12 18-APR-1990;
FEATURES
source Location/Qualifiers
1..583
/organism="unknown"

BASE COUNT 188 a 83 c 93 g 219 t
 ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 583;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 381 TTCTCAGAA 389
 |||||
 Cp 33 ttctcagaa 25

RESULT 29
 LOCUS A48440 585 bp DNA PAT 07-MAR-1997
 DEFINITION Sequence 5 from Patent WO9602641.
 ACCESSION A48440
 NID 92302229
 VERSION A48440.1 GI:2302229
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

REFERENCE 1 (bases 1 to 585)
 COOPER,C.S. and Gusterson,B.A.
 MATERIALS AND METHODS RELATING TO THE DIAGNOSIS AND PROPHYLACTIC
 AND THERAPEUTIC TREATMENT OF SYNOVIAL SARCOMA
 Patent: WO 9602641-A 5 01-FEB-1996;
 JOURNAL CANCER RES CAMPAIGN TECH (GB)
 COMMENT OTHER Publication AD 2986595 960216.
 FEATURES
 source
 1..585
 /organism="unidentified"
 /db_xref="taxon:32644"
 1..477
 /note="unamed protein product. Protein sequence is in
 conflict with the conceptual translation"
 /codon_start=1
 /protein_id="CAA03105.1"
 /db_xref="PID:e306311"
 /db_xref="PID:92302230"
 /db_xref="GI:2302230"
 /translation="QQDDAYGGPPPGQGYPPQGGYPCQGGYSGGSGPG
 QYPPYPOGGQGGYGGRTQPPQPPQGGYGYDQIMPKKPADENDSKGVSASGP
 QNDGKQLHPKPKANISEKINRSKPKRKHMTLRERKQIVYIEIISDEEDDE"

BASE COUNT 186 a 155 c 156 g 88 t
 ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 585;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 351 TTCTCAGAA 359
 |||||
 Oy 25 ttctcagaa 33

RESULT 30
 LOCUS I74318 662 bp DNA PAT 23-DEC-1997
 DEFINITION Sequence 16 from patent US 5688641.
 ACCESSION I74318
 NID 93010459
 VERSION I74318.1 GI:3010459
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 662)
 SAGER,R., ZOU,Z., Lee,S.Whan and Tomasetto,C.Laure.
 TITLE Cancer diagnosis using nucleic acid hybridization
 JOURNAL Patent: US 5688641-A 16 18-NOV-1997;
 FEATURES
 source
 1..662
 /organism="unknown"

BASE COUNT 179 a 158 c 195 g 130 t
 ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 662;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 460 TTCTCAGAA 468
 |||||
 Cp 33 ttctcagaa 25

RESULT 31
 LOCUS AB023802S2 668 bp DNA VRT 04-MAR-1999
 DEFINITION Gallus gallus Rd gene for riboflavin binding protein, exon 6 and
 partial cds.
 ACCESSION AB023803
 NID 94512597
 VERSION AB023803.1 GI:4512597
 KEYWORDS Rd; riboflavin binding protein.
 SEGMENT 2 of 2
 SOURCE Gallus gallus adult male liver DNA, clone_lib:chicken genomic
 library.

ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
 Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (sites)
 Zheng,D.B., Lim,H.M., Pene,J.J. and White,H.B. III.
 Chicken riboflavin-binding protein. cDNA sequence and homology with
 milk folate-binding protein
 J. Biol. Chem. 263 (23), 11126-11129 (1988)
 88298752
 JOURNAL 2 (bases 1 to 668)
 MEDLINE
 HamaJima,S.
 DIRECT Submission
 Submitted (17-FEB-1999) to the DDBJ/EMBL/Genbank databases. Susumu
 HamaJima, Iwate Medical University, Biochemistry; 19-1 Uchimarui,
 Morioka, Iwate 020-8505, Japan (E-mail:shamaj@iwate-med.ac.jp,
 Tel:81-19-651-5111(ex.3274), Fax:81-19-653-9244)
 FEATURES
 source
 1..668
 /organism="Gallus gallus"
 /note="Rd mRNA consists of exons in AB023244, X74246,
 AB023802 and AB023803"
 /db_xref="taxon:9031"
 /clone_lib="chicken genomic library"
 /dex_stage="adult"
 /sex="male"
 /tissue_type="liver"
 /join(AB023802.1:1..4377,1..668)
 /gene="Rd"
 /join(AB023802.1:<1..129,AB023802.1:3542..3653,312..548)
 /gene="Rd"
 /codon_start=2
 /product="riboflavin binding protein"
 /protein_id="BA075356.1"
 /db_xref="PID:d1039094"
 /db_xref="PID:94512599"
 /db_xref="GI:4512599"
 /translation="CEDFTKIKCEFTKCSFHARWIDPRTAIOYVPLCOSFCDDMY
 EACKDSDICAHNMLTDWERSGSEHNCKSKCVSEYANGTMCOSMGESKRVSES
 SCICLQNNKKDMVAIKHLSESSSESSSMSSSEHACOKKLLKFEALQDEGEERR"
 <1..311
 /gene="Rd"
 /number=5
 312..668
 /gene="Rd"
 /number=6
 /product="riboflavin binding protein"
 645..650
 /gene="Rd"
 668
 /gene="Rd"
 polyA_site
 polyA_signal
 exon
 intron

BASE COUNT 194 a 135 c 162 g 177 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 668;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 633 TTCGAGAA 641
|||||
QY 25 ttctcagaa 33

RESULT 32
LOCUS SCMGUPPER 701 bp mRNA INV 30-JUN-1994
DEFINITION Schistosoma mansoni glutathione peroxidase mRNA, complete cds.
ACCESSION M86510
NID 9160997
VERSION M86510.1 GI:160997
KEYWORDS
SOURCE Schistosoma mansoni (strain Puerto Rican) adult cDNA to mRNA.
ORGANISM Schistosoma mansoni
Eukaryotes; mitochondrial eukaryotes; Metazoa; Platyhelminthes;
Trematoda; Digenea; Strigeidida; Schistosomatoidea;
Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 701)
AUTHORS Williams, D.L., Pierce, R., Cookson, E. and Capron, A.
TITLE Molecular cloning and sequencing of glutathione peroxidase from Schistosoma mansoni
JOURNAL Mol. Biochem. Parasitol. 52, 127-130 (1992)
FEATURES 92326859
Location/Qualifiers
1. 701
/organism="Schistosoma mansoni"
/strain="Puerto Rican"
/db_xref="taxon:6183"
/dev_stage="adult"
1. 15
/EC_number="1.11.1.9"
/note="selenocysteine"
/codon_start=1
/trans_except=(pos:142..144,aa:Cys)
/product="glutathione peroxidase"
/protein_id="AA29885.1"
/db_xref="PID:g160998"
/db_xref="GI:160998"
/translation="MSSSHKSNISYEFYTKDINGVSLKRYGHVCLIVNACG
ATDKNRYQDQEMTRLVGKGLRIAPCCNPGCEPAPAEIKKPTLTKYGVQPMES
KIRVNSDADDLTKFLKSRQHGILINNINNESKFLVDRGQGVKRSPTTADYDIES
DIMELEK"
526. 701
3'UTR polyA_signal 682.. 687
BASE COUNT 213 a 129 c 147 g 212 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 701;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 675 TTCGAGAA 683
|||||
Cp 33 ttctcagaa 25

RESULT 33
LOCUS MSGCGAB5B 738 bp DNA BCT 08-JAN-1998
DEFINITION Mycobacterium bovis gene for alpha antigen, antigen 85B, MPB59,
antigen 6, partial cds.
ACCESSION D78142
NID 91060915
VERSION D78142.1 GI:1060915
KEYWORDS MPB59; alpha antigen; antigen 6; antigen 85B.

SOURCE Mycobacterium bovis (strain:BCG Tokyo) DNA.
ORGANISM Mycobacterium bovis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE 1 (bases 1 to 738)
AUTHORS Ohara, N.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1995) to the DDBJ/EMBL/GenBank databases. Naoya
Ohara, Nagasaki University School of Dentistry, Oral Bacteriology;
Sakamoto, I.-T., Nagasaki, Nagasaki 852, Japan
(E-mail: f02028cc.nagasaki-u.ac.jp, tel: 0958-49-7649,
Fax: 0958-49-7650)
2 (bases 492 to 738)
REFERENCE Matsuo, K., Yamaguchi, R., Yamazaki, A., Tsasaka, H. and Yamada, T.
AUTHORS Cloning and expression of the Mycobacterium bovis BCG gene for
TITLE extracellular alpha antigen
JOURNAL J. Bacteriol. 170 (9), 3847-3854 (1988)
MEDLINE 88314872
JOURNAL 3 (sites)
AUTHORS Ohara, N., Nishiyama, T., Ohara-Wada, N., Matsumoto, S., Matsuo, T. and
Yamada, T.
TITLE Characterization of the transcriptional initiation regions of genes
for the major secreted protein antigens 85C and MPB51 of
JOURNAL Mycobacterium bovis BCG
MEDLINE Microb. Pathog. 23 (5), 303-310 (1997)
FEATURES 98070721
Location/Qualifiers
1. 738
/organism="Mycobacterium bovis"
/strain="BCG Tokyo"
/db_xref="taxon:1765"
509. 514
/citation=[2]
533. 538
/citation=[2]
569. 573
/citation=[2]
583. >738
/citation=[2]
/codon_start=1
/transl_table=11
/product="alpha antigen, antigen 85B, MPB59, antigen 6"
/protein_id="BAA11219.1"
/db_xref="PID:d1011881"
/db_xref="PID:g1060916"
/db_xref="GI:1060916"
/translation="MTDVSRRKIRAGRRMLMISTAAVYLPGLVGLAGAAVAGASRP
GIPVETIQ"
583. 702
sig_peptide 703. >738
mat_peptide
BASE COUNT 140 a 237 c 226 g 135 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 738;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 395 TTCGAGAA 403
|||||
QY 25 ttctcagaa 33

RESULT 34
LOCUS TAU32225 768 bp DNA INV 01-AUG-1995
DEFINITION Theileria annulata cytochrome oxidase polypeptide III (coxIII) .
gene, mitochondrial gene encoding mitochondrial protein, complete
cds.
ACCESSION U32225
NID 915547
VERSION U32225.1 GI:915547
KEYWORDS

SOURCE Theileria annulata.
Mitochondrion Theileria annulata
Eukaryotes; mitochondrial eukaryotes; Alveolata; Apicomplexa;
Piroplasmida; Theileridae; Theileria.

REFERENCE 1 (bases 1 to 768)
AUTHORS Lawson, D. and Hall, R.
TITLE Synonymous codon usage (SCU) bias differentiates Theileria nuclear
genes from mitochondrial ones
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 768)
AUTHORS Lawson, D.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1995) Daniel Lawson, Department of Biology,
University of York, P.O. Box 373, York YO1 5YW, UK

FEATURES
Source
1..768
/organism="Theileria annulata"
/mitochondrion
/strain="Hissar"
/db_xref="taxon:5874"
/dev_stage="piroplasm"
1..768
/gene="coxIII"
1..768
/gene="coxIII"
/EC_number="1.9.3.1"
/note="structural protein of the electron transport chain"
/codon_start=1
/transl_table=4
/function="electron transport"
/product="cytochrome oxidase polypeptide III"
/protein_id="AA073631.1"
/db_xref="PID:9915548"
/db_xref="GI:915548"
/translation="MRSVQSDLRKININISERLYFGTDLLEYIDSTYKRFIL
MYNQHLVGTLYKLVSGEFFINSLTIFINGIKETMTSTVIMATIGMTFSEIMV
FSTFGFFHRLSPWMIIEVNEAFQISDVNAGSILSLIQRIEERGFEVDY
MLERILIGFIFLSQGDYSLVKSYNNHWTLYFNLTGLSHLYVVGIFALMQA
FASNGCGOKEDFNAGYMHVEIIMTALMLFL"
BASE COUNT 266 a 84 c 116 g 302 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 768;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 289 TTCTCAGAA 297
|||||
Cp 33 ttctcagaa 25

RESULT 35
LOCUS CHU34232 789 bp mRNA MAM 26-MAR-1999
DEFINITION Capra hircus interferon-gamma mRNA, complete cds.
ACCESSION U34232
NID 91730273
VERSION U34232.1 GI:1730273
KEYWORDS
SOURCE goat.
ORGANISM Capra hircus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
Caprinae; Capra.
1 (bases 1 to 789)
AUTHORS Beyer, J.C., Stich, R.W., Hoover, D.S., Brown, W.C. and Cheevers, W.P.
TITLE Cloning and expression of caprine interferon-gamma
JOURNAL Gene 210 (1), 103-108 (1998)
MEDLINE 98192545
REFERENCE 2 (bases 1 to 789)
AUTHORS Beyer, J.C. and Cheevers, W.P.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) Joseph C. Beyer, Veterinary Microbiology
and Pathology, Washington State University, Bustad Hall, Pullman,

FEATURES
Source
1..789
/organism="Capra hircus"
/db_xref="taxon:9925"
/cell_type="peripheral blood mononuclear cells"
56..556
/codon_start=1
/product="interferon-gamma"
/protein_id="AA938525.1"
/db_xref="PID:91730274"
/db_xref="GI:1730274"
/translation="MKYSSFLALLSVLGFSGSGYCGPFKEIKENLKEYFNASND
VAKGPIFSEILKMKESDRIKIOSQVSEFYFLFNKDNQVIRSDMIKQDMFO
KFLNGSEKLEDFKKLIQIPVDDIQIRKAINELIKYMNLSPRKRRKRRSGLNR
GRRSW"
BASE COUNT 260 a 141 c 146 g 242 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 789;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 209 TTCTCAGAA 217
|||||
Cp 33 ttctcagaa 25

RESULT 36
LOCUS BS16SRNS 808 bp DNA BCT 04-SEP-1991
DEFINITION B. subtilis 16S rRNA gene, 5' end.
ACCESSION X55332
NID 939779
VERSION X55332.1 GI:39779
KEYWORDS 16S ribosomal RNA.
SOURCE Bacillus subtilis.
ORGANISM Bacillus subtilis
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;
Bacillus.
1 (bases 1 to 808)
AUTHORS Jarvis, E.D., Cheng, S. and Rudner, R.
TITLE Genetic structure and DNA sequences at junctions involved in the
rearrangements of Bacillus subtilis strains carrying the trpE26
mutation
JOURNAL Genetics 126 (4), 785-797 (1990)
MEDLINE 91169259
FEATURES
Source
1..808
/organism="Bacillus subtilis"
/strain="168-type"
/db_xref="taxon:1423"
/clone="pWR305"
1..476
/note="ORF"
/codon_start=3
/transl_table=1
/protein_id="CA39031.1"
/db_xref="PID:9388214"
/db_xref="GI:388214"
/db_xref="SWISS-PROT:O04385"
/translation="AFFEANOFLDAYNAVVAHLHRLARIEVIDRGFHPETVMSQVRO
MEPVYKLYSELISHSRLRLPLFLHSAEIGSAHLEPMKEKDIQMPG
ELLQHDHLKFTQDLGVMYLTLEKGLINVCQIETKQAVYHHRGYSFKKGVDS"
460..465
/note="p1 promoter"
483..488
/note="p1 promoter"
557..562
/note="p2 promoter"
580..585
/note="p2 promoter"
767..>808
/note="rind ribosomal operon"

BASE COUNT 261 a 162 c 174 g 211 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 808;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 508 TTCTGAGAA 516
|||||
Cy 25 ttctgagaa 33

RESULT 37
LOCUS AFAJ3315 855 bp mRNA PLN 27-FEB-1998
DEFINITION Aspergillus fumigatus mRNA for rasp f 7 allergen.
ACCESSION AJ223315
NID 92879887
VERSION AJ223315.1 GI:2879887
KEYWORDS allergen; rasp f 7.
SOURCE Aspergillus fumigatus.
ORGANISM Aspergillus fumigatus.
Eukaryota; Fungi; Ascomycota; Euascomycetes; Plectomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 855)
AUTHORS Hemmann, S.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Hemmann S., Molecular Allergy, Swiss
Institute of Allergy and Asthma Research (SIAR), Obere Strasse 22,
CH-7270 Davos, SWITZERLAND
2 (bases 1 to 855)
REFERENCE Cramer, R.
TITLE Recombinant Aspergillus fumigatus allergens: from the nucleotide
sequences to clinical applications
JOURNAL Int. Arch. Allergy Immunol. 115 (2), 99-114 (1998)
MEDLINE 96141802
FEATURES
source
1. 855
/organism="Aspergillus fumigatus"
/strain="ATCC 4202"
/db_xref="taxon:5085"
/clone_lid="Filamentous phage display cDNA library"
1. 339
/gene="rasp f 7"
/gene="rasp f 7"
/gene="rasp f 7"
/codon_start=1
/product="rasp f 7"
/protein_id="CAA11255.1"
/db_xref="PID:e1250610"
/db_xref="PID:g2879888"
/db_xref="GI:2879888"
/db_xref="SPRMBL:042799"
/translation="SSGSGPCSGSCVGGOLYYDNTASAPSSGCLTNDGSENV
VALPVGIMTADGCKTIVTITNGITKATAYVDKCMGCKPDLDAKRLBELDFPSAG
RIDGMSWTFN"
BASE COUNT 213 a 230 c 196 g 216 t
ORIGIN
Query Match 100.0%; Score 9; DB 27; Length 855;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 TTCTGAGAA 128
|||||
Cy 25 ttctgagaa 33

RESULT 38
LOCUS A25392 890 bp DNA PAT 03-MAR-1995
DEFINITION Oligonucleotide.
ACCESSION A25392
NID 9633562

VERSION A25392.1 GI:833582
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 890)
AUTHORS
TITLE METHOD FOR SELECTING RECOMBINANT MICRO-ORGANISMS OF WHICH THE
SURFACE COMPRISES AT LEAST ONE MOLECULE HAVING ENZYMATIC ACTIVITY
JOURNAL Patent: WO 9311242-A 6 10-JUN-1993;
FEATURES
source
1. 890
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 236 a 206 c 228 g 220 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 890;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 303 TTCTGAGAA 311
|||||
Cy 33 ttctgagaa 25

RESULT 39
LOCUS CHKREP 900 bp mRNA VRT 30-SEP-1988
DEFINITION Chicken riboflavin-binding protein mRNA, complete cds.
ACCESSION J03922
NID 9212630
VERSION J03922.1 GI:212630
KEYWORDS riboflavin-binding protein.
SOURCE Chicken (White leghorn, adult female) oviduct, cDNA to mRNA.
ORGANISM Gallus gallus
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 900)
AUTHORS Zheng, D.B., Lam, H.M., Pene, J.J. and White, H.B.III.
TITLE Chicken riboflavin-binding protein: cDNA sequence and homology with
milk folate-binding protein
JOURNAL J. Biol. Chem. 263, 11126-11129 (1988)
MEDLINE 88296752
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by H.B. White III, 24-JUN-1988.
FEATURES
source
1. 900
/organism="Gallus gallus"
/db_xref="taxon:9031"
64. 105
/note="riboflavin-binding protein signal peptide"
64. 780
/note="riboflavin-binding protein precursor"
/codon_start=1
/protein_id="AAA49056.1"
/db_xref="PID:g212631"
/db_xref="GI:212631"
/translation="MLRFAITLFAVITSTCOYGLGEGDTHKAPSPPEPNHHECTLY
SSSSCCYANFTBOLASPIIKVSNYSWNGGOLSKSCDEFTKIECFYAGSCHAARMI
DRYRTAIGSVPLCOSEFCDDWYKACKDSDICAHNMTDPERDSGHNHCKSKCPYSE
MTANGIDMCOSMGSEFKVSESSCICLONKKDMAIKHLSESSSESSSSSEBHA
CCKLKFELQOESEERK"
106. 777
/note="riboflavin-binding protein"
BASE COUNT 268 a 196 c 231 g 205 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 900;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 865 TTCTGAGAA 873

|||||||
 25 ttctcagaa 33

RESULT 40 TETPRAN 933 bp mRNA INV 05-FEB-1999
 LOCUS Tetrachymena pyriformis mRNA for Ran/TC4, complete cds.
 DEFINITION
 ACCESSION D21825
 D21825
 MID G532502
 VERSION D21825.1 GI:532502
 KEYWORDS GTP-binding low-Mr protein; Ran/TC4.
 SOURCE Tetrachymena pyriformis (strain:W) cDNA to mRNA, clone_lib:lambda
 gtl0.

ORGANISM Tetrachymena pyriformis
 Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
 Hymenostomatida; Tetrachymenina; Tetrachymena.
 REFERENCE 1 (bases 1 to 933)
 AUTHORS Nagata,K.
 JOURNAL Direct Submission
 Submitted (27-OCT-1993) to the DDBJ/EMBL/GenBank databases.
 Koh-ichi Nagata, Gifu University School of Medicine, Department of
 Biochemistry, 40 Tsukasamachi, Gifu 500, Japan
 (Tel:0582-65-1241(ex.2230), Fax:0582-65-9002)
 2 (bases 1 to 933)
 REFERENCE 2 (bases 1 to 933)
 AUTHORS Nagata,K., Takemasa,T., Alam,S., Hattori,T., Watanabe,Y. and
 Nozawa,Y.
 JOURNAL Cloning of cDNAs encoding a cell-cycle-regulatory GTP-binding
 low-Mr (GBLM) protein, Ran/TC4, from micronucleated Tetrachymena
 thermophila and amicronucleated Tetrachymena pyriformis
 Gene 144 (1), 123-125 (1994)
 MEDLINE 94299154
 COMMENT Submitted (27-Oct-1993) to DDBJ by:
 Koh-ichi Nagata
 Department of Biochemistry
 Gifu University School of Medicine
 40 Tsukasamachi
 Gifu 500
 Japan
 Phone: 0582-65-1241 x2230
 Fax: 0582-65-9002.

FEATURES

Source
 1..933
 /organism="Tetrachymena pyriformis"
 /strain="W"
 /db_xref="taxon:5908"
 /clone_lib="lambda gtl0"
 38..709
 /codon_start=1
 /transl_table=6
 /product="Ran/TC4"
 /protein_id="BAA04849.1"
 /db_xref="PID:d1005384"
 /db_xref="PID:9559385"
 /db_xref="GI:559385"
 /translation="MVDNKQDYVAEFLKLVLDGSGVGTFTFVTHQGTGEFFKRYIARQ
 GVNATNMLHTKGAIRFNIMDTAGQETLGRGYTGADAAIMEDVSRITKNI
 PKMKDLTRICENPIVYGNKVDSDKSVARQITFRKRSLOYYDVSASKNYER
 PFWILRLKIDPMLNVEGIALAPVDIHMTEDQIKQLQMHDMNNAQCGLPDBE
 DDEFN"
 CDS
 301 a 220 c 170 g 242 t

BASE COUNT 301 a 220 c 170 g 242 t
 ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 933;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 531 TTCACGAA 539
 |||||
 Cp 33 ttctcagaa 25

RESULT 41 BTRGCE03 963 bp DNA MAN 04-FEB-1999
 LOCUS

DEFINITION Bos taurus membrane guanylate cyclase isoform E precursor (GC-E)
 gene, exons 5, 6 and 7.
 ACCESSION U77098
 U77098
 MID G2182818
 VERSION U77098.1 GI:2182818
 KEYWORDS guanylate cyclase; retina; bovine GC-E gene; photoreceptor.
 SEGMENT 3 of 9
 SOURCE Bos taurus.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 963)
 AUTHORS Johnston,J.P., Farhangfar,F., Aparicio,J.G., Nam,S.H. and
 Applebury,M.L.
 JOURNAL The bovine guanylate cyclase GC-E gene and 5' flanking region
 Gene 193 (2), 219-227 (1997)
 MEDLINE 97398155
 REFERENCE 2 (bases 1 to 963)
 AUTHORS Johnston,J.P., Farhangfar,F., Aparicio,J.G., Nam,S.H. and
 Applebury,M.L.
 JOURNAL Direct Submission
 Submitted (04-NOV-1996) The Howe Laboratory, Harvard Medical School
 /MEI, 243 Charles St., Boston, MA 02114, USA
 /Location/Qualifiers
 1..963

FEATURES

Source
 1..963
 /organism="Bos taurus"
 /strain="Holstein dairy cow"
 /db_xref="taxon:9913"
 /clone_lib="Stratagene #3605"
 /tissue_type="Retina"
 <1..8
 /gene="GC-E"
 /number=4
 9..93
 /gene="GC-E"
 /number=5
 94..362
 /gene="GC-E"
 /number=5
 363..465
 /gene="GC-E"
 /number=6
 466..851
 /gene="GC-E"
 /number=6
 852..953
 /gene="GC-E"
 /number=7
 954..>963
 /gene="GC-E"
 /number=7
 Intron
 199 a 278 c 255 g 231 t
 BASE COUNT 199 a 278 c 255 g 231 t
 ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 963;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 769 TTCACGAA 777
 |||||
 Cp 33 ttctcagaa 25

RESULT 42 AF047171 1001 bp DNA BCT 19-MAR-1998
 LOCUS Providencia stuartii plasmid VR-1 beta-lactamase TEM-60 (Dia
 DEFINITION
 ACCESSION AF047171
 AF047171
 MID G2970344
 VERSION AF047171.1 GI:2970344
 KEYWORDS Providencia stuartii.
 SOURCE Providencia stuartii
 ORGANISM

Eubacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Providencia

REFERENCE 1 (bases 1 to 1001)
 AUTHORS Franceschini, N., Perilli, M., Segatore, B., Setacci, D., Amicosante, G., Mazzariol, A. and Cornaglia, G.
 TITLE Ceftazidime and aztreonam resistance in *Providencia stuartii*: characterization of a natural TEM-derived ESBL

JOURNAL 2 (bases 1 to 1001)
 REFERENCE Franceschini, N., Perilli, M. and Amicosante, G.
 AUTHORS Direct Submission
 TITLE Submitted (09-FEB-1998) Sciences and Biomedical Technologies, University of L'Aquila, Via Vetoio - Coppito, L'Aquila, AQ 67100, Italy

FEATURES
 source location/Qualifiers
 1..1001 /organism="Providencia stuartii"
 /plasmid="VR-1"
 /db_xref="taxon:588"
 136..996 /gene="bla TEM-60"
 136..996 /gene="bla TEM-60"
 /ec_number="3.5.2.6"
 /note="beta-lactam hydrolase; natural TEM-derived extended spectrum beta lactamase (ESBL)"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase TEM-60"
 /protein_id="AAC05975.1"
 /db_xref="PID:g2970345"
 /db_xref="GI:2970345"
 /translation="MSIOHFRVALIPFPAFCPLVFAHPETLVKVKDAEDQLGSVGY IELDPSGKLTSEFRRPEPRPMSTFKVLGAVLRVDAGQOLGRIRIHVSNDLVK YSPVTEKHLTDGMTVREICSAITMSDNTAANLLTTGGSOELTAFIHNMGDHYTRL DSWPEINERAIPIVNDERDTTPRAMRTTKRLTLGELLTLASROQLIDMEADVAGPL LRCALPAGMLIDKSGGERSGRTIAALGPDKPSRIYVITTSQATMDERNROIA EIGASLIKHM"

BASE COUNT 271 a 227 c 249 g 254 t

ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 1001;
 Best Local Similarity 100.0%; Pred. No. 8,17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 420 TTCTCAGAA 428
 |||||||
 Cp 33 ttctcagaa 25

RESULT 43
 LOCUS SWTEMA0GE 1011 bp DNA BCT 25-NOV-1997
 DEFINITION S.marcescens beta-lactamase gene.
 ACCESSION X97254
 MID 92648042
 VERSION X97254.1 GI:2648042
 KEYWORDS beta-lactamase.
 SOURCE Serratia marcescens.
 ORGANISM Serratia marcescens
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Serratia

REFERENCE 1 (bases 1 to 1011)
 PERILLI, M., FELICI, A., FRANCESCHINI, N., DE SANTIS, A., PAGANI, L., LUZZARO, F., ORATORE, A., ROSSOLINI, G. M., KNOX, J. R. and AMICOSANTE, G.
 TITLE Characterization of a new TEM-derived beta-lactamase produced in a *Serratia marcescens* strain

JOURNAL Antimicrob. Agents Chemother. 41 (11), 2374-2382 (1997)
 MEDLINE 98037387
 REFERENCE 2 (bases 1 to 1011)
 PERILLI, M.
 AUTHORS Direct Submission
 TITLE Submitted (12-APR-1996) M. Perilli, Dipartimento di Scienze e Technologie Biomediche e Biometria, Universita dell'Aquila, Via

Vetoio, Loc. Coppito, 69100 L'Aquila, ITALY
 REMARK revised by submitter 03-JUN-96

FEATURES
 source location/Qualifiers
 1..1011 /organism="Serratia marcescens"
 /strain="S5"
 /db_xref="taxon:615"
 145..148 /note="putative"
 154..1011 /ec_number="3.5.2.6"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="CAA65888.1"
 /db_xref="PID:el187395"
 /db_xref="PID:g2648043"
 /db_xref="GI:2648043"
 /db_xref="SPTREMBL:O33677"
 /translation="MSIOHFRVALIPFPAFCPLVFAHPETLVKVKDAEDQLGSVGY IELDPSGKLTSEFRRPEPRPMSTFKVLGAVLRVDAGQOLGRIRIHVSNDLVK YSPVTEKHLTDGMTVREICSAITMSDNTAANLLTTGGSOELTAFIHNMGDHYTRL DSWPEINERAIPIVNDERDTTPRAMRTTKRLTLGELLTLASROQLIDMEADVAGPL LRCALPAGMLIDKSGGERSGRTIAALGPDKPSRIYVITTSQATMDERNROIAE IGASLIKHM"

BASE COUNT 270 a 231 c 245 g 265 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 1011;
 Best Local Similarity 100.0%; Pred. No. 8,17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 438 TTCTCAGAA 446
 |||||||
 Cp 33 ttctcagaa 25

RESULT 44
 LOCUS AF031352 1047 bp DNA INV 01-SEP-1998
 DEFINITION Euplotes aediculatus ER heat shock protein 70 (HSP70) gene, partial cds.
 ACCESSION AF031352
 MID 93169832
 VERSION AF031352.1 GI:3169832
 KEYWORDS
 SOURCE Euplotes aediculatus.
 ORGANISM Euplotes aediculatus.
 Eukaryote; Alveolata; Ciliophora; hypotrichs; Euplotida; Euplotes.

REFERENCE 1 (bases 1 to 1047)
 BUDIN, K. and PHILIPPE, H.
 TITLE new insights into the phylogeny of eukaryotes based on ciliate HSP70 sequences

JOURNAL Mol. Biol. Evol. 15 (8), 943-956 (1998)
 MEDLINE 98384834
 REFERENCE 2 (bases 1 to 1047)
 BUDIN, K.
 AUTHORS Direct Submission
 TITLE Submitted (24-OCT-1997) Laboratoire de Biologie Cellulaire, Universite Paris Sud, Bat. 444, Orsay 91405, France

FEATURES
 source location/Qualifiers
 1..1047 /organism="Euplotes aediculatus"
 /db_xref="taxon:5940"
 /note="endoplasmic reticulum"
 <1..>1047 /gene="HSP70"
 /product="heat shock protein 70"
 <1..>1047 /gene="HSP70"
 <1..>1047 /gene="HSP70"
 /codon_start=1
 /transl_table=10

```

/product="heat shock protein 70"
/protein_id="AAC33417.1"
/db_xref="PID:93169833"
/db_xref="GI:3169833"
/translation="TSCVGIKNGRVEIIPNELGNRTTPSFVAFTEDERLVGEAAKN
QAALNPERSIYVIRKLIGNFNEKQVDKLVSYNIYDKNGKPYVSAVNGELKVM
SPEISAMILVMKQIAENYVGOEVKNVITVPAYFNDAORATADAGIISGLNARI
LNEPTAAIAVGLDKKPGEOITLFDLGGGFVSLTIDNEVFEEVETAGTHLGG
DEDOVLOVFKIKIPKYNKOISTDKRAIOKKGSKAKRSLSSVIOKTEIEVVE
GVDEKETITRAFEELCADLFRKKTLEPVEQALDKGKKKTDIDELVVGSTRIRKVO
OLIKDFNGKERNRG"
BASE COUNT      370 a      195 c      232 g      250 t
ORIGIN
Query Match      100.0%; Score 9; DB 22; Length 1047;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      470 TTCTGAGAA 478
|||||||
OY      25 ttctgagaa 33

RESULT 45
LOCUS      AF051814      1050 bp      DNA      BCT      02-JAN-1999
DEFINITION Escherichia coli strain ECOR 52 major pilin protein precursor
            (papa) and paph (paph) genes, partial cds.
ACCESSION  AF051814
NID         94106020
VERSION     AF051814.1 GI:4106020
KEYWORDS
SOURCE
ORGANISM   Escherichia coli.
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE  1 (bases 1 to 1050)
AUTHORS   Boyd,E.F. and Hatli,D.L.
TITLE     Diversifying selection governs sequence polymorphism in the major
            adhesin proteins FimA, Papa, and SfaA in Escherichia coli
JOURNAL    J. Mol. Evol. (1999) In press
REFERENCE  2 (bases 1 to 1050)
AUTHORS   Boyd,E.F.
TITLE     Direct Submission
JOURNAL    Submitted (03-MAR-1998) Department of Organismic & Evolutionary
            Biology, Harvard University, 16 Divinity Ave., Cambridge, MA 02138
FEATURES
            source
            1..1050
            /organism="Escherichia coli"
            /strain="ECOR 52"
            /db_xref="taxon:562"
            <1..525
            /gene="papa"
            <1..525
            /gene="papa"
            /codon_start=1
            /product="papa"
            /translation="PMAYVSEGVNAPNIPGOGKVTENGTVVADPCISIKSADOSI
            DFGQSKFLEAGVSKPMDDIEVNCIDITFKGNSAGOKGTVALATGPIVNHSD
            ELDITNGTGTALVVOGAKNVVFDGSEGDANTLKGEVNLHTYAVVKKSSAVGAAYTE
            GAFSAVANPNLTLYO"
            589..>1050
            /gene="paph"
            589..>1050
            /gene="paph"
            /codon_start=1
            /product="paph"
            /translation="PMAYVSEGVNAPNIPGOGKVTENGTVVADPCISIKSADOSI
            DFGQSKFLEAGVSKPMDDIEVNCIDITFKGNSAGOKGTVALATGPIVNHSD
            ELDITNGTGTALVVOGAKNVVFDGSEGDANTLKGEVNLHTYAVVKKSSAVGAAYTE
            GAFSAVANPNLTLYO"
            /db_xref="PID:94106021"
            /db_xref="GI:4106021"
            /db_xref="PID:94106022"
            /db_xref="GI:4106022"
            /protein_id="AAD02665.1"
            /db_xref="PID:94106022"
            /db_xref="PID:94106022"

```

```

/db_xref="GI:4106022"
/translation="MRLFSVLPFFCGVHGVFAGFPFPPGKSLPEYGEHHWMD
GRAFHGVVBPACTLAMEDAMWOIIDGETPVRRLONGFSSEPKFSLRPNCFNSO
GGNFSDSRIRIVTIDGVAGETPDPEFNLGSKAKGINTLOIGDARGNIARAAGVM"
BASE COUNT      254 a      203 c      286 g      307 t
ORIGIN
Query Match      100.0%; Score 9; DB 18; Length 1050;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      105 TTCTCAGAA 113
|||||||
CP      33 ttctcagaa 25

RESULT 46
LOCUS      KPY17583      1071 bp      DNA      BCT      29-JUN-1998
DEFINITION Klebsiella pneumoniae blaTEM-22 gene, partial.
ACCESSION  Y17583
NID         93282027
VERSION     Y17583.1 GI:3282027
KEYWORDS   beta-lactamase class A.
SOURCE     Klebsiella pneumoniae.
ORGANISM   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Klebsiella.
REFERENCE  1 (bases 1 to 1071)
AUTHORS   Arlet,G., Goussard,S., Courvalin,P. and Philippou,A.
TITLE     Sequence of the genes for TEM-20, TEM-21, TEM-22 and TEM-29
            extended-spectrum beta-lactamases
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1071)
AUTHORS   Arlet,G.J.
TITLE     Direct Submission
JOURNAL    Submitted (15-JUN-1998) G.J. Arlet, Hopital Saint-Louis
            Microbiologie, 1 Avenue Claude Vellefaux, 75475 Paris Cedex 10,
            FRANCE
FEATURES
            source
            1..1071
            /organism="Klebsiella pneumoniae"
            /strain="SLK52"
            /db_xref="taxon:573"
            144..148
            167..172
            214..1071
            /gene="blaTEM-22"
            214..>1071
            /gene="blaTEM-22"
            /codon_start=1
            /product="beta-lactamase class A"
            /translation="MSIOHFRVALIPFEAFCLPVFAHPETLVVKADEKLGARVY
            IELDNGSKILESFRPERPMPMSTFVLLCGAVLSRYDAGDQEDLRRIHYSONDIYK
            YSPYTERKHLTGKRTVRELCSAITMSDNTANLLITIGSKRELTAFLHNNGDVYTR
            DRWPELNEALPNDERDITTPAAVATYLRKLTLBELTLASRQQLIDWEADKVAGPL
            LRSALPAGWFIADKSGSGSERSGRIITIALGPDGKPRIVVITYTGSQATMDERRNOIA
            EIGASLILKHW"
            293 a      234 c      263 g      281 t
ORIGIN
Query Match      100.0%; Score 9; DB 17; Length 1071;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      498 TTCTCAGAA 506
|||||||
CP      33 ttctcagaa 25

```

RESULT 47
LOCUS SUB131631 1079 bp DNA BCT 11-MAR-1999
DEFINITION Streptococcus uberis, skc gene, strain 9756-296-LF (SK889).
ACCESSION A131631
NID 94033580
VERSION A131631.1 GI:4033580
KEYWORDS skc gene; streptokinase.
SOURCE Streptococcus uberis.
ORGANISM Streptococcus uberis.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 1079)
AUTHORS Johnson, L.B.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1998) Johnson L.B., Molecular and structural
biology, University of Aarhus, Gustav Wiedsvej 10C, 8000 C,
DENMARK
REFERENCE 2 (bases 1 to 1079)
AUTHORS Johnson, L.B., Poulsen, K., Kilian, M. and Petersen, T.E.
TITLE Purification and cloning of a streptokinase from Streptococcus
uberis
JOURNAL Infect. Immun. 67 (3), 1072-1078 (1999)
MEDLINE 99150235
FEATURES
source location/Qualifiers
1. 1079
/organism="Streptococcus uberis"
/isolate="9756-296-LF"
/db_xref="taxon:1349"
181..255
/gene="skc"
181..1041
/gene="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
/protein_id="CA10462.1"
/db_xref="PID:e1360474"
/db_xref="PID:g4033581"
/db_xref="GI:4033581"
/translation="MKKFLIMLIGFGCATOPSKVAITGDSDYARYIDPDKK
ITPAINDGVEGSSNOEILRGITHVITDONKIVTAELLDAIRHOMVLODYSYE
LVDPADPAOLTRDRRLLFANRNEESVEDTIOEYLKGHVILKRVVEPTTHPE
TANLEYVVOGATKDGEPHPLIFVDYGEKIGETLDEPKIAEKLQSYDYMID
QKEXTIILKNSLGLQIPRYVSPDFSYEIODRIMAKDKPSGHELGSETSIDNVEK
YLTPKSYKP"
181..1041
/gene="skc"
/gene="skc"
256..1038
/gene="skc"
/product="streptokinase"
BASE COUNT 384 a 173 c 174 g 348 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 1079;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 617 TTCTGAGAA 625
|||||
Cp 33 ttctcagaa 25

RESULT 48
LOCUS 131987 1100 bp DNA PAT 07-JAN-1997
DEFINITION Sequence 2 from patent US 5585099.
ACCESSION 131987
NID 91822778
VERSION 131987.1 GI:1822778
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 1100)
AUTHORS Richards, S., Kaplan, J. and Mosciacki, R.
TITLE Prolectin as a vaccine adjuvant
JOURNAL Patent: US 5585099-A 2 17-DEC-1996.
FEATURES
source location/Qualifiers
1. 1100
/organism="unknown"
BASE COUNT 324 a 282 c 202 g 292 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1100;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 977 TTCTGAGAA 985
|||||
Cy 25 ttctcagaa 33

RESULT 49
LOCUS 117239 1106 bp DNA PAT 24-JAN-1996
DEFINITION Sequence 4 from patent US 5486463.
ACCESSION 117239
NID 91252147
VERSION 117239.1 GI:1252147
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1106)
AUTHORS Lesslauer, W., Lotscher, H. and Stuber, D.
TITLE TNF-mutins
JOURNAL Patent: US 5486463-A 4 23-JAN-1996.
FEATURES
source location/Qualifiers
1. 1106
/organism="unknown"
BASE COUNT 305 a 256 c 247 g 298 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1106;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 379 TTCTGAGAA 387
|||||
Cy 25 ttctcagaa 33

RESULT 50
ID E08994 standard; DNA; UNC; 1110 bp.
AC E08994;
SV E08994.1
NI d1107331
DT 07-OCT-1997 (Rel. 52, Created)
DE 07-OCT-1997 (Rel. 52, Last updated, Version 1)
DE DNA encoding an immunogenic peptide derived from human Influenza A
DE virus haemagglutinin that doesn't have globular region.
KW JP 1995089992-A/4.
OS unidentified
OC unclassified.
RN [1]
RP 1-1110
RA Okuno Y., Isekawa Y., Sasao F., Ueda S.;
RT "IMMUNOGENIC ARTIFICIAL POLYPEPTIDE";
RL Patent number JP 1995089992-A/4, 04-APR-1995.
RL TAKARA SHUZO CO LTD.
CC OS None
CC OC Artificial sequences.
CC PN JP 1995089992-A/4
CC PD 04-APR-1995
CC PF 16-MAR-1994 JP 1994070194
CC PR 20-APR-1993 JP 93P 115216
CC PI OKUNO YOSHINOBU, ISEKAWA YUJI, SASAO FUYOKO, UEDA SHIGEHARU

CC PC C07K7/08.A61K39/145.C12N15/44.C12P21/02//C12N15/06.C12P21/08,
CC PC (C12P21/02,
CC PC C12R1:91),(C12P21/08.C12R1:91).C07K99:00;
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC FH Key Location/Qualifiers
CC FT source 1. 1110
CC FT /organism="Artificial sequences"
CC FT CDS 34..1077
CC FT /product="immunogenic peptide"
CC FT sig-peptide 34..81
FH Key Location/Qualifiers
FT source 1. 1110
FT /organism="unidentified"
FT /db_xref="taxon:32644"
SQ Sequence 1110 BP: 374 A; 216 C; 258 G; 262 T; 0 other;
Query Match 100.0%; Score 9; DB 11; Length 1110;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 619 TTCTCAGAA 627
|||||
Cp 33 ttctcagaa 25
RESULT 51
LOCUS CEL236896 1134 bp mRNA INV 08-MAR-1999
DEFINITION Caenorhabditis elegans mRNA for GRP1 protein.
ACCESSION AJ236896
VERSION g4481794
NID AJ236896.1 GI:4481794
KEYWORDS GRP1 protein;
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia;
Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 1134)
AUTHORS Venkateswarlu, K.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1999) Venkateswarlu K., Biochemistry, The
University of Bristol, School of Medical Sciences, University Walk,
BS8 1JD, UNITED KINGDOM
2 (bases 1 to 1134)
REFERENCE 2 (bases 1 to 1134)
AUTHORS Venkateswarlu, K. and Cullen, P.J.
TITLE Cloning and functional characterisation of C. elegans GRP1
JOURNAL Unpublished
FEATURES
source
1. 1134
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
1. 1134
/gene="K06H7.4"
1. 1134
/gene="K06H7.4"
/function="GTP-exchange factor for ADP-ribosylation
factors"
/note="general receptor for phosphoinositides-1"
/codon_start=1
/product="GRP1 protein"
/protein_id="CAB38534.1"
/db_xref="pid:e1392686"
/db_xref="pid:g4481795"
/db_xref="gi:4481795"
/translation="MTLPKVRKKAQLVDEIEALKEVREYDEELDQVYYTHPKSKY
HKIVNGKKEFNODPKALDWLASRNVAKPOALALRMKAGEGLSKAIGELIDNR
PEALETLDRTKEHKLHDVPIYPAIRVLFSPRPGESOKINRIEKAELYANQNP
YGNADQATVAKSCIMVNTLHNPVYKRPSEKIEINEDOLEGATITQUTVYE
SVSVQFKIPDEVSTSGKIVNDILHAERGWLKKSSNPLFSGALSWMKRWPLSE
NCLVYFDKMTKEPGITLLANVGIRKVEASRPMPFIFSLSDQIACRKTEDGRL
VEGRHSYKICAVNDEDMRSMWNAISRMAVQOHLARPKSTH"

BASE COUNT 401 a 211 c 239 g 283 t
ORIGIN
Query Match 100.0%; Score 9; DB 21; Length 1134;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 822 TTCTCAGAA 830
|||||
Cp 25 ttctcagaa 33
RESULT 52
LOCUS AF053028 1140 bp DNA MAM 31-MAR-1998
DEFINITION Panthera tigris altaica isolate S3 cytochrome b (cytb) gene,
mitochondrial gene encoding mitochondrial protein, complete cds.
ACCESSION AF053028
VERSION g2995746
NID AF053028.1 GI:2995746
KEYWORDS
SOURCE tiger.
ORGANISM Mitochondrion Panthera tigris
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Carnivora; Fissipedia; Felidae; Panthera.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Cracraft, J., Feinstein, J., Vaughn, J. and Helm-Bychowski, K.
TITLE Sorting out tigers (Panthera tigris): mitochondrial sequences,
nuclear inserts, systematics, and conservation genetics
JOURNAL Anim. Conserv. 1, 63-74 (1998)
2 (bases 1 to 1140)
REFERENCE Cracraft, J., Feinstein, J., Vaughn, J. and Helm-Bychowski, K.
AUTHORS Direct Submission
JOURNAL Submitted (10-MAR-1998) Ornithology, American Museum of Natural
History, 79th at Central Park West, New York, NY 10024, USA
FEATURES
source
1. 1140
/organism="Panthera tigris"
/mitochondrion
/sub_species="altaica"
/isolate="S3"
/db_xref="taxon:9694"
1. 1140
/gene="cytb"
1. 1140
/gene="cytb"
/codon_start=1
/translation="cytochrome b"
/product="cytochrome b"
/protein_id="AAC08547.1"
/db_xref="pid:g2995747"
/db_xref="gi:2995747"
/translation="MTNIRKSHPIIKIINHSFIDLPAPSNISAMNFGSLGVCILIO
ILTGLFLAMHTSPTMTAFSSVTHICHDVNRGQIRIRLHANGAMFICLYMNRGRM
YIGSTTSETWNIIGVILFVMAFAGVYLPWQMSFWATVITNLISATPIGTDL
VEMWIGSDVKATLRFPAHFILPFSVSLAVALHLEFHTSSNNSGVSDSDKI
PFHPYTIKIDILGLVLLITLILFSPDLGPDVYIPANLPNTPHILKPEWYFIF
AYATRSIPNKGVALVLSILITARIIPALHTSKORGMFRPISOCFLWLVADLLT
LTWIGQVPEHPFALIGOLASILTEFILLVMPISGIENRLMK"

ACCESSION region.
 NID 036198
 VERSION 91223870
 KEYWORDS 039198.1 GI:1223870
 SEGMENT 1 of 8
 SOURCE plg.
 ORGANISM Sus scrofa
 Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Artiodactyla; Suidae; Sus.
 REFERENCE 1 (bases 1 to 1204)
 AUTHORS Szafranska, B., Xie, S., Miura, R. and Roberts, R.M.
 TITLE Direct Submission
 JOURNAL Submitted (23-OCT-1995) Bozena Szafranska, Animal Sciences, University of Missouri-Columbia, 158 ASRC, Columbia, MO 65211, USA
 COMMENT Szafranska B. Biol. Reprod 53, 21-28, 1995.
 FEATURES
 source Location/Qualifiers
 1..1204
 /organism="Sus scrofa"
 /note="5' flanking region of PPAG2"
 /db_xref="taxon:9823"
 /rissue_type="placenta"
 BASE COUNT 325 a 309 c 252 g 318 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 23; Length 1204;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 442 TTCGACAA 450
 ||||||||
 Qy 25 ttctgaga 33
 RESULT 54
 LOCUS A04190 1212 bp DNA PAT 16-APR-1993
 DEFINITION Artificial sequence for galactosidase fusion protein containing P.knowlesi CS protein immunogenic region.
 ACCESSION A04190
 NID 9344796
 VERSION A04190.1 GI:344796
 KEYWORDS fusion protein.
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 1212)
 AUTHORS
 TITLE PROTECTIVE PEPTIDE ANTIGEN
 JOURNAL Patent: WO 8402922-A 7 02-AUG-1984;
 FEATURES
 source Location/Qualifiers
 1..1212
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..1212
 /codon_start=1
 /transl_table=1
 /product="galactosidase fusion protein"
 /protein_id="CAA00346.1"
 /db_xref="PID:9344797"
 /db_xref="GI:344797"
 /translation="MSIQHFRVALIPFFAFCLEVFRAHPTLYKVKDAEDQAGRVY
 IELDLSNGKILLESFRPEERPMSTFVLLCGAVLSVHDAGQDQGRRIHYSQNDLVE
 YSPVTEKHLTDGMTRELCSAATITMSDNTANLLITIGPKELTAFLLHMGDHYTL
 DRMEPELNEAIPNDERDTMPAGGGGGGAGAGOPQAGDAGANAGOPQAGDAN
 AGOPQAGDAGANAGOPQAGDAGANAGOPQAGDAGANAGOPQAGDAGANAGOPQAGD
 ANAGOPQAGDAGANAGOPQAGDAGANAGOPQAGDAGANAGOPQAGDAGANAGOPQAGD
 LNSALPAGWFIADKSGAGERSRGITIALGPDGKPSRIYVITITGSOATWDERNROI
 AEIGASLIRKM"
 BASE COUNT 351 a 287 c 337 g 237 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 1212;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 285 TTCGACAA 293
 ||||||||
 Cp 33 ttctgaga 25
 RESULT 55
 LOCUS RAU59717 1234 bp DNA BCT 27-JUN-1996
 DEFINITION Rickettsia akari citrate synthase (gltA) gene, partial cds.
 ACCESSION U59717
 NID 91389976
 VERSION U59717.1 GI:1389976
 KEYWORDS
 SOURCE
 ORGANISM
 Rickettsia akari.
 Eubacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 Rickettsiaceae; Rickettsiae; Rickettsia.
 REFERENCE 1 (bases 1 to 1234)
 AUTHORS Raoult, D., Rydkina, E., Roux, V., Eremeeva, M. and Balayeva, N.
 TITLE Phylogenetic Analysis of the genus Rickettsia by Citrate Synthase gene Sequencing
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1234)
 AUTHORS Raoult, D.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-1996) Faculte de Medecine, Unite des Rickettsies, CNRS EP 00054, 27, bd Jean Moulin, 13385 Marseille Cedex 5, France
 FEATURES
 source Location/Qualifiers
 1..1234
 /organism="Rickettsia akari"
 /strain="MK (Kaplan)"
 /db_xref="taxon:786"
 1..1234
 /gene="gltA"
 <1..>1234
 /gene="gltA"
 /codon_start=2
 /transl_table=11
 /product="citrate synthase"
 /protein_id="FAB02953.1"
 /db_xref="PID:91389977"
 /db_xref="GI:1389977"
 /translation="DSFAELKIRGKIFOLPLIKASIGEDVIDISRVSAEVCFTYDP
 GFNSTASCKSTITTYIDGKILRHRYNFKDLAERSDFLVAVLLYGLPSIEDYNN
 FTKVVAHSLVNERLHYLFQPTCSSHPMAIMLAISAFYPLILNKEDYEITA
 IRIRAKIPITPIAMSYKYSIGQPEVYDPSNLDFTENLHMFAPPCRYKVNIIKNAL
 NKFTLIHADHONASTSTRVIAAGSGANAFACISGIVASIAMGAPAGANEAIVIMLEK
 IGSENIPIRYIAKAKSDSPRLMGCHRYVKNYDRAVATKCTCEVKEKELGQLEN
 PLQINILELALAKDEYIERKLPVNDVFGIITYKAGIPSOFTVLFALARTVGM
 MAQWKEMHEDPOKISRPR"
 BASE COUNT 410 a 207 c 236 g 381 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 17; Length 1234;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 838 TTCGACAA 846
 ||||||||
 Qy 25 ttctgaga 33
 RESULT 56
 LOCUS TCR6226 1241 bp DNA INV 15-FEB-1999
 DEFINITION Trypanosoma cruzi mpx gene encoding mitochondrial peroxiredoxin.
 ACCESSION AJ006226
 NID 93393045
 VERSION AJ006226.1 GI:3393045
 KEYWORDS antioxidant; mpx gene; peroxiredoxin.
 SOURCE Trypanosoma cruzi.
 ORGANISM Trypanosoma cruzi.
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE
AUTHORS
TITLE
JOURNAL

Trypanosoma; Schizotrypanum.
1 (bases 1 to 1241)
Wilkinson, S. R.

Direct Submission

Submitted (18-MAY-1998) Wilkinson S. R., Department of Infectious Diseases, London School of Hygiene and Tropical Medicine, Pathogen Molecular Biology and Biochemistry Unit, Keppel Street, London, WC1E 7HT, UK

REFERENCE
AUTHORS
JOURNAL

2 (bases 1 to 1241)
Wilkinson, S. R., Temperton, N. J., Mondragon, A. and Kelly, J. M.
Unpublished

FEATURES

Location/Qualifiers

1..1241

/organism="Trypanosoma cruzi"

/strain="CL-Brener"

/db_xref="taxon:5693"

/dev_stage="epimastigote"

/gene="mpx"

274..954

/gene="mpx"

/codon_start=1

/evidence="experimental"

/product="peroxiredoxin"

/protein_id="CA06923.1"

/db_xref="PID:e1387021"

/db_xref="PID:94388655"

/db_xref="GI:4388655"

/translation="MFRMAVTSIQKLSRAFCNTLRNLNDYQAKTKATVRPAE
WAGCAVNGKIKIDISLNDYKGVLLFYPADFFVCPTEITAFSAQAEKINTV
VAASCDQSHLAWINTPRNKGIGEMSIPTLSQTEIKEDYGVLLIEQGISRGIF
IIDKGLRHRTVNDLPVGRNVEVLNVQAFQYVDKNKDVIPTCMRPGKPTMTKEKA
NEVEKNA"

274..954

/gene="mpx"

385..951

/gene="mpx"

/product="peroxiredoxin"

BASE COUNT 341 a 235 c 277 g 388 t

ORIGIN

gene

mat_peptide

Query Match 100.0%; Score 9; DB 21; Length 1241;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1128 TTCTCAGAA 1136

Cp 33 ttctcagaa 25

RESULT 57
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM

1253 bp DNA BCT 11-DEC-1997
P. furiosus cpkA gene.
Y09829
GI:2370353
carbamoyl kinase: carbamoyl-phosphate synthetase; cpkA gene.
Pyrococcus furiosus.
Pyrococcus furiosus.
Archaea; Euryarchaeota; Thermococcales; Thermococcaceae;

Pyrococcus.

1 (bases 1 to 1253)

Durbecq, V., Legrain, C., Roovers, M., Pierard, A. and Giansdorff, N.
The carbamate kinase-like carbamoyl phosphate synthetase of the
hyperthermophilic archaeon Pyrococcus furiosus, a missing link in
the evolution of carbamoyl phosphate biosynthesis
Proc. Natl. Acad. Sci. U.S.A. 94 (24), 12803-12808 (1997)

JOURNAL

MEDLINE

98058725

2 (bases 1 to 1253)

Durbecq, V.

Direct Submission

Submitted (05-DEC-1996) V. Durbecq, Laboratoire de Microbiologie
UB and Institut de recherches du CERIA, 1 Av. Emile Grison, 1070

JOURNAL

AUTHORS

TITLE

REFERENCE

AUTHORS

JOURNAL

MEDLINE

98058725

2 (bases 1 to 1253)

Durbecq, V.

Direct Submission

Submitted (05-DEC-1996) V. Durbecq, Laboratoire de Microbiologie
UB and Institut de recherches du CERIA, 1 Av. Emile Grison, 1070

JOURNAL

AUTHORS

TITLE

REFERENCE

AUTHORS

JOURNAL

MEDLINE

98058725

REMARK
COMMENT
On Sep 9, 1997 this sequence version replaced g1:1743371.
Location/Qualifiers

source

1..1253

/organism="Pyrococcus furiosus"

/strain="DSM 3638"

/db_xref="taxon:2261"

79..84

/note="Box A"

108..112

/note="Box B"

118..1062

/gene="cpkA"

118..1062

/gene="cpkA"

/codon_start=1

/transl_table=11

/product="carbamate kinase-like carbamoylphosphate
synthetase"

/protein_id="CAA70972.1"

/db_xref="PID:e286171"

/db_xref="PID:91743372"

/db_xref="GI:1743372"

/db_xref="SPTREMBL:P95474"

/translation="MGRKRVIALGNAALQGGKSGYEEMDNVRKTARQIAEIIAR
KEYVITHNGNGPOVGSILHMDAGQATYGIPOADPDYAGASOGNIGIMQALKEIR
KRMKERRVYITIGOTIVDKNDPAFONPTKVGPFDEFTAKRLARRGKWRKDSRG
WRKRVSPSPDKGVEAEFTIKKLYERGVYIVASGGGVPTVLEGEIKGVAVYDKDLA
GEEALAEVNDIMILLTDVNGALVYGTETEDWLREKVEELKRYEEGHFKGSMGP
KVLAARFIMGGERIALIAHLEKVAVALAEKGTOTVLP"

1065..1073

terminator

BASE COUNT 394 a 214 c 337 g 308 t

ORIGIN

gene

Query Match 100.0%; Score 9; DB 17; Length 1253;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 TTCTCAGAA 34

Cp 33 ttctcagaa 25

RESULT 58
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM

1297 bp mRNA MAM 01-MAR-1999
C. familliaris mRNA for T-cell surface glycoprotein CD4.
X68565
GI:288652
X68565.1 GI:288652
HIV binding; MHC class II molecule binding; T cell surface
glycoprotein; transmembrane protein.
dog.
Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 1297)

Malde, K. F.

Direct Submission

Submitted (05-OCT-1992) K. F. Malde, University of Miami School,
Diabetes Research Institute (R-134), PO Box 016960, Miami FL 33101,
USA

JOURNAL

MEDLINE

93192324

2 (bases 1 to 1297)

Malde, K. F., Conner, G. E., Mintz, D. H. and Alejandro, R.

Primary structure of the canine CD4 antigen

Biochim. Biophys. Acta 1172 (3), 315-318 (1993)

JOURNAL

MEDLINE

93192324

Location/Qualifiers

1..1297

/organism="Canis familiaris"

/strain="Beagle"

/db_xref="taxon:9615"

/tissue_type="thymus"

University of Oxford, Dept of Biochemistry, South Parks Road,
Oxford OX1 3OU, UK

FEATURES
Source Location/Qualifiers
1..1305
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_1ib="lambda gtl1 bovine liver cDNA"
sig_peptide 148..207
CDS 148..1263
/codon_start=1
/product="conglutinin"
/protein_id="CA50665.1"
/db_xref="PID:Q395268"
/db_xref="GI:395268"
/db_xref="SWISS-PROT:P23805"
/translation="MLPLPLSYLLLLTTPMNSLGAEMTTFPSOKLIANACTLYMCSPL
SGLPGRHDGQDGECHPEKGRGDPGSPGAPGRGPRGVPPIGKGDNGFVGEPRGDT
GPRGPGMPGAPGREGSPGKQSGMGPPTGPKGTGKRGVAGPGLGFPSPGLK
EKAGPGETGAPGAGVTPGSAIGQSGAPGAPGLKGRDGPETGAKGSGLA
NALKORVTLIDGLRFRFONAFSQRKAVLFFPGCAVGEKIFETAGAVKSYSDVLCR
EAKGOLASPRSAEMEAAYTOMVAOENKAYLSMNDISTEGRTYPTGELIYVSNMAD
EPNNSDEQPEKCVETFPDGKNDVPCSKQLVCEP"
mat_peptide 208..1260
/product="conglutinin"
BASE COUNT 322 a 340 c 405 g 238 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 1305;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 225 TTCTGAGAA 233
|||||
Cp 33 ttctcgaaa 25

RESULT 61
LOCUS 109391 1310 bp PAT 14-NOV-1994
DEFINITION Sequence 24 from Patent WO 8906703.
ACCESSION 109391
NTD 9587921
VERSION 109391.1 GI:587921
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1310)
AUTHORS Dryja,T.P., Friend,S. and Yandell,D.W.
DIAGNOSIS OF RETINOBLASTOMA
JOURNAL Patent: WO 8906703-A 24 27-JUL-1989;
FEATURES Location/Qualifiers
1..1310
source /organism="unknown"
BASE COUNT 464 a 192 c 219 g 435 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1310;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 195 TTCTGAGAA 203
|||||
Qy 25 ttctcgaaa 33

RESULT 62
LOCUS AV277768 1346 bp DNA PLN 08-JAN-1997
DEFINITION A.velutinum chloroplast rpoa gene.
ACCESSION 277768
NTD 91769911
VERSION 277768.1 GI:1769911
KEYWORDS pcd gene; ribosomal protein 11; RNA polymerase; rpoa gene; rps11
gene.

SOURCE
ORGANISM Australopyrum velutinum.
Chloroplast Australopyrum velutinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Australopyrum.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Petersen,G. and Seberg,O.
TITLE Phylogenetic analysis of the Triticeae (Poaceae) based on rpoa
sequence data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1346)
AUTHORS Petersen,G.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1996) G. Petersen, Botanical Institute, Botanical
Laboratory, Gothersgade 140, DK-1123 Copenhagen K, DENMARK
FEATURES Location/Qualifiers
1..1346
/organism="Australopyrum velutinum"
/chloroplast
/isolate="H6724"
/db_xref="taxon:58935"
/dev_stage="adult"
/tissue_type="leaves"
1..24
/gene="rps11"
<1..24
/codon_start=1
/product="ribosomal protein 11"
/protein_id="CA01382.1"
/db_xref="PID:e256588"
/db_xref="PID:Q1769913"
/db_xref="GI:1769913"
/translation="PPKKRR"
96..1115
/gene="rpoa"
96..1115
/gene="rps11"
/codon_start=1
/product="alpha subunit of RNA polymerase"
/protein_id="CA01381.1"
/db_xref="PID:e256852"
/db_xref="PID:Q1769912"
/db_xref="GI:1769912"
/db_xref="SPRMBL:P93993"
/translation="MVRREVAGSTOTLOMKCVESRVDSKRLYYGRFLLSPLRKGAUT
VGIALRALGEITGCTITRAKRSVPHSEYTTIGIESVQETLLNKEIYLRNLTG
VRDASICYKGRPTTADITLIPSYEIVDTROPANLTPEDICIDQIKRDRQTE
LRKNYQGSYPIIDAVXMPVRNVNSISCGGNKHEILLEIWTNGLTPKEALYEA
SRNIDILFELPLAAEEGTSFEENKRNFTPLFTFORLTLNKKNGGIPNCIFIDQ
LELTSRYNCIKRANIHITLIDLKSTEDLRLIDSFMEDRKHIWDTLEKHLPIDLK
NKLSE"
complement(1320..1346)
/gene="pchl"
complement(1320..>1346)
/gene="pchl"
/codon_start=1
/protein_id="CA01383.1"
/db_xref="PID:e256589"
/db_xref="PID:Q1769914"
/db_xref="GI:1769914"
/db_xref="SPRMBL:P93992"
/translation="KSLTLGLF"
BASE COUNT 472 a 220 c 248 g 404 t 2 others
ORIGIN

Query Match 100.0%; Score 9; DB 27; Length 1346;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1012 TTCTGAGAA 1020
|||||
Qy 25 ttctcgaaa 33

```

RESULT 63
LOCUS MSE223384 1346 bp mRNA INV 12-MAR-1998
DEFINITION Manduca sexta mRNA for 26S proteasome regulatory ATPase subunit 10b
            (S10b).
ACCESSION AJ223384
NID 92960215
VERSION AJ223384.1 GI:2960215
KEYWORDS ATPase.
SOURCE tobacco hornworm.
ORGANISM Manduca sexta
          Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
          Pterygota; Lepidoptera; Sphingioidea; Sphingidae; Sphinginae;
          Manduca.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Hastings, R.A.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Hastings R.A., School of Biomedical
          Sciences, University of Nottingham, Medical School, Queen's Medical
          Centre, Nottingham, Notts, NG7 2UH, UK
2 (bases 1 to 1346)
REFERENCE Hastings, R.A., Dawson, S.P., Billelt, M.A. and Mayer, R.J.
AUTHORS The regulatory ATPases S10b is not only present in the 26S
TITLE proteasome but also in a smaller 220 kDa complex and as free
          molecules in the intersegmental muscles of the tobacco horn moth,
          Manduca sexta
JOURNAL Unpublished
FEATURES
SOURCE Location/Qualifiers
1. .1346
   /organism="Manduca sexta"
   /db_xref="taxon:7130"
90. .1280
   /gene="Ms.S10b"
90. .1280
   /gene="Ms.S10b"
   /product="26S proteasome regulatory ATPase subunit 10b
   (S10b)"
   /protein_id="CA11285.1"
   /db_xref="PID:e1266343"
   /db_xref="PID:g2860216"
   /db_xref="GI:2960216"
   /db_xref="SPTREMBL:C62556"
   /translation="MPAGSAVDTPPRKARQDYRKIKMEKEVESRLKAMRDQLKLT
   KOYKSENDIKALQSGOIVGEVVKOLEEETFKATNGRPRVYVCGRRDKNKLG
   TRVALDMTTLTIRHLRREVDPIVYNNMSHEDPGPYTAAGIGLOEIOLEIOLREYELPL
   LNPFLRYRVGCTTPPKGLILGPGTGTTLARASOLDANFLKVSASAIYDKYIGS
   ARLTREMTNARDHQPCITIFMDEIDAGGRFSSTADREIQTLLELLNOMDGFDS
   LGQVKIIMATNRPTLDPALRGLRLKLEIPLPNQARLEILKIHAAPIAHGEMD
   YEAVVGLSDTFENGADLRNVCTEAGLFIARAEREYIIODIMLKAVRYADMRKLESKLD
   YKPY"
BASE COUNT 402 a 311 c 349 g 284 t
ORIGIN
Query Match 100.0%; Score 9; DB 21; Length 1346;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

          Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;
          Secernentea; Rhabdilia; Rhabditida; Rhabditina; Rhabditioidea;
          Rhabditidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 1365)
AUTHORS Reisenstein, K.M. and Emmons, S.W.
TITLE Nematode repetitive DNA with ARS and segregation function in
          Saccharomyces cerevisiae
JOURNAL Mol. Cell. Biol. 8, 875-883 (1988)
MEDLINE 88174732
FEATURES
SOURCE Location/Qualifiers
1. .1365
   /organism="Caenorhabditis elegans"
   /db_xref="taxon:6239"
BASE COUNT 497 a 249 c 281 g 333 t
ORIGIN
Query Match 100.0%; Score 9; DB 21; Length 1360;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 113 TTCACGAA 121
CP 33 ttctcagaa 25
RESULT 65
LOCUS AF106032 1365 bp DNA BCT 15-MAR-1999
DEFINITION Bacillus stearothermophilus replicative helicase (dnab) gene,
          complete cds.
ACCESSION AF106032
NID 94416321
VERSION AF106032.1 GI:4416321
KEYWORDS
SOURCE Bacillus stearothermophilus.
          Bacillus stearothermophilus.
          Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;
          Bacillus.
REFERENCE 1 (bases 1 to 1365)
AUTHORS Bird, L.E. and Wigley, D.B.
TITLE The Bacillus stearothermophilus replicative helicase: cloning,
          overexpression and activity
JOURNAL Unpublished
2 (bases 1 to 1365)
REFERENCE Bird, L.E. and Wigley, D.B.
AUTHORS Direct Submission
JOURNAL Submitted (11-NOV-1998) Sir William Dunn School of Pathology,
          University of Oxford, South Parks Road, Oxford OX1 3RE, UK
FEATURES
SOURCE Location/Qualifiers
1. .1365
   /organism="Bacillus stearothermophilus"
   /strain="NCA1503"
   /db_xref="taxon:1422"
1. .1365
   /gene="dnab"
1. .1365
   /gene="dnab"
   /note="5'-3' helicase; similar to Bacillus subtilis Dnab;
          Dnab"
   /codon_start=1
   /transl_table=11
   /product="replicative helicase"
   /protein_id="FAD20314.1"
   /db_xref="PID:g4416322"
   /db_xref="GI:4416322"
   /translation="MSLFSERIPPOSIEAEQAVLGAVFLDPAALVPASEILIPEDFY
   RAHQKIFHAMLKRVADRGEPVDLVYTAELASQLEEGVSYSELADAVPTANV
   EYARIVKESVIRLIRPATSIADQGYRDEIDVLLEADRKIMEVQRHSGAFK
   NIDILVQTDNTEMLHNRDGEITGIPTEFELDRTSFGQSDILIVAAKPSGTA
   FALNINQNVATKTNENVALFSLEMSAQQLVPRALCEGNNQNLRTGLTPEDWCKL
   TVMAGSLSNAGIYIDTPSIRVSDIRAKCRKQESGLGIVIDVLQILQSGSKEN
   ROOEVEISRSKLARELEVPVIALSQLSRSVEORODRPMQMSIRREGSGTEQADAI
   VALIYRDVYNNQSENKNIIEIIIAKQNRGPGVQLATIKKYNNFVNLERFDEAOI
   PRGA"

```

BASE COUNT 349 a 346 c 378 g 292 t
ORIGIN
Query Match 100.0%; Score 9; DB 18; Length 1365;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1215 TTCTGAGAA 1223
|||||
25 ttctgagaa 33
RESULT 66
LOCUS CEU10327 1368 bp mRNA INV 01-APR-1995
DEFINITION Caenorhabditis elegans MAD homolog 1 (cem-1) mRNA, complete cds.
ACCESSION U10327
NID 9551486
VERSION U10327.1 GI:551486
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdillia; Rhabdillina; Rhabdilloidea;
Rhabdillidae; Caenorhabditis.
REFERENCE
AUTHORS
1 (bases 1 to 1368)
Seklisky,J.J., Newfield,S.J., Raftery,L.A., Charloff,E.H. and
Gelbart,W.M.
TITLE
Genetic characterization and cloning of mothers against dpp, a gene
required for decapentaplegic function in Drosophila melanogaster
JOURNAL
MEDLINE
Genetics 139 (3), 1347-1358 (1995)
REFERENCE
AUTHORS
2 (bases 1 to 1368)
Seklisky,J.J.
TITLE
Mothers against dpp: A gene required for decapentaplegic function
in Drosophila melanogaster
JOURNAL
Thesis (1993) Cellular and Developmental Biology, Harvard
University
3 (bases 1 to 1368)
Gelbart,W.M.
TITLE
Direct Submission
JOURNAL
Submitted (02-JUN-1994) William M. Gelbart, Cellular and
Developmental Biology, Harvard University, Biological Laboratories,
16 Divinity Avenue, Cambridge, MA 02138, USA
FEATURES
source
Location/Qualifiers
1..1368
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone_lib="mixed stage cDNA library (Barstead and
Waterston, 1989)"
/chromosome="III"
45..1301
/gene="cem-1"
45..1301
/gene="cem-1"
/note="cem-1"
/note="Similar to Drosophila melanogaster MAD polypeptide,
GenBank Accession Number U10328"
/codon_start=1
/product="MAD homolog 1"
/protein_id="AAC46583.1"
/db_xref="PID:9551487"
/db_xref="PID:9551487"
/translation="MINDGKIKITIRLTKMGDEENMARKAIDNMLKLIKHNKQA
LENLEFALRCGQOCTECVTPRSLDGLQISHRALPHVYICRYRWPDLSHHEK
AIEDRCFESGQDICIINPYKRVKATGLPVLVRYSEKPEOEVPTLAKELM
EMSGRMPQNNMNVNMTANOFHOYNGNGIEEMDTSKRPIDPGVPCVAPPKVE
EQFNATVSYELNTRVGEQVAVYSTTTITIDFTPTPCINGSLISGLSNVRNATIE
TRRHGNVSKLTVKNSGLTAQCSDSATVVOSSNCTYINGFSTTVKLANCSLK
IFDMEIFRQLLEDCSRGFDASFDLQKMTFLRFSFKGWAQYORODVTSPCWIEIH
LHAPLAMDRLVLTSGTPPRPISSTIS"
BASE COUNT 457 a 256 c 273 g 382 t
ORIGIN
Query Match 100.0%; Score 9; DB 22; Length 1368;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 632 TTCTCAGAA 640
|||||
33 ttctcagaa 25
RESULT 67
LOCUS DREHOXD4 1400 bp mRNA VRT 02-DEC-1997
DEFINITION Danio rerio mRNA for Hoxd4 protein.
ACCESSION Y14548
NID 92661826
VERSION Y14548.1 GI:2661826
KEYWORDS Hoxd4 protein.
SOURCE
Zebrafish.
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinoidae; Cyprinidae; Rasbora; Danio.
REFERENCE
AUTHORS
1 (bases 1 to 1400)
Prince,V.E.
TITLE
Direct Submission
JOURNAL
Submitted (08-AUG-1997) V.E. Prince, Princeton University,
Molecular Biology, Washington Road, Princeton, NJ 08544, USA
REFERENCE
AUTHORS
2 (bases 1 to 1400)
Prince,V.E., Joly,L., Ekker,M. and Ho,R.K.
TITLE
Zebrafish hox genes: genomic organization and modified collinear
expression patterns in the trunk
JOURNAL
Unpublished
FEATURES
source
Location/Qualifiers
1..1400
/organism="Danio rerio"
/db_xref="taxon:7955"
/dev_stage="embryo"
369..1139
/gene="hoxd4"
369..1139
/gene="hoxd4"
/function="transcription factor"
/codon_start=1
/product="Hoxd4 protein"
/protein_id="CAA4863.1"
/db_xref="PID:e1198500"
/db_xref="PID:92661827"
/db_xref="GI:2661827"
/db_xref="SPRMBL:057374"
/translation="MEGKKDKNSKISISITLQIKLMSSTVNSKYVDFPCEETS
ONSTIPQSPFYSPSQDTDFQHPGITSRSNYSQPTICSTVGSSVQPRGHVQDAS
TPSPFPQOTOCQPAVVOISGSRTCGQOONTONGIPKOPAVVYPMWKKVHYTVND
YTGPEPKRSRTATROOVLELEKEKHFNRRLTRRRRIESHATLSLSRQIKWIFQNR
MKMKDKDKLNTGSRASVGNHGHQKQSGREITFL"
BASE COUNT 440 a 310 c 297 g 353 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 1400;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 497 TTCTCAGAA 505
|||||
33 ttctcagaa 25
RESULT 68
LOCUS A32415 1412 bp mRNA PAT 13-JUN-1996
DEFINITION H. sapiens PAI-2 cDNA in pDBPL.
ACCESSION A32415
NID 91567408
VERSION A32415.1 GI:1567408
KEYWORDS
SOURCE
human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1412)

REFERENCE 1
AUTHORS POLYPEPTIDE PRODUCTION IN FUNGI
TITLE Patent: WO 9102057-A 19 21-FEB-1991;
JOURNAL Location/Qualifiers
FEATURES
1..1412
source /organism="Homo sapiens"
/db_xref="taxon:9606"

gene 42..1289
/gene="PAI-2"
42..1289
/gene="PAI-2"
/codon_start=1
/product="plasma1ogen activator inhibitor 2"
/protein_id="CAA02099.1"
/db_xref="PID:e249479"
/db_xref="PID:g1567409"
/db_xref="GI:1567409"
/translation="MEDICANTLFLALNFKHLAKASPNQNLSPMSISSTMAVYM
GSRGSTDQAKVLOENEGCANVTPPTPTSCGFMQIOKSGSPDALQQAADK
IHSSFRSLSSAINSTGNTLVESYNKLEKSSAFREIYRLQKITSSEQAADLE
CAEARKKINSWKTQTKGIPLNLEPGSDGTQVNLVAVFKGKMKPEFKLNG
LYPFRVSAQRTPVQMYLREKLNIGYIEDLKQIIEPYAGVSMFLPLDIAVVS
TGLELSEITDYKLNKMTSKDKMAEDEEVYIPOFKLEEHYELRSILSRMGEDAFN
KGRANFGMSERNDLFLSEVFOHVMVDNEGEAAAGTGVMTGRGHGQPGFVADH
PFLIMHKTKITNIFGRGRSSP"

BASE COUNT 418 a 311 c 323 g 360 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1412;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1009 TTCTGACA 1017
|||||||
QY 25 ttctgaga 33

RESULT 69
LOCUS AGTTT576 1429 bp DNA PLN 03-MAY-1995
DEFINITION A.githago L. nrDNA (Oxelman ITS-AGR30616 (GB)).
ACCESSION X86895
NID 9798816
VERSION X86895.1 GI:798816
KEYWORDS 28S ribosomal RNA; 5.8S ribosomal RNA; internal transcribed spacer
1; internal transcribed spacer 2; ITS1; ITS2.
SOURCE
organism Agrostemma githago
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Caryophyllales; Caryophyllales; Caryophyllaceae; Agrostemma.
1 (bases 1 to 1429)
Oxelman,B. and Liden,M.
Generic boundaries in the tribe Sileneae (Caryophyllaceae) as
inferred by nuclear rDNA sequences
Unpublished
REFERENCE 2 (bases 1 to 1429)
Oxelman,B.
Direct Submission
Submitted (03-MAY-1995) Oxelman B., Department of Systematic
Botany, Carl Skottsbergs Gata 22, S-413 19, Gothenburg, SWEDEN
3 (bases 1 to 1429)
Oxelman,B. and Liden,M.
Generic boundaries in the tribe Sileneae (Caryophyllaceae) as
inferred by nuclear rDNA sequences
Taxon 44, 525-542 (1995)
Location/Qualifiers
1..1429

JOURNAL /organism="Agrostemma githago"
FEATURES /note="Species author is L."
- source

/isolate="voucher specimen Oxelman ITS-AGR30616 (GB)"
/db_xref="taxon:39848"
/issue_type="fresh leaf"
1..226
/note="Internal transcribed spacer 1"
227..381
/product="5.8S ribosomal RNA"
382..615
/note="Internal transcribed spacer 2"
616..>1429
/product="28S ribosomal RNA"
391 c 459 g 269 t

BASE COUNT 310 a 391 c 459 g 269 t
ORIGIN

Query Match 100.0%; Score 9; DB 27; Length 1429;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1368 TTCTGACA 1376
|||||||
QY 25 ttctgaga 33

RESULT 70
LOCUS ECOCOLE2IC 1457 bp DNA BCT 07-FEB-1999
DEFINITION Escherichia coli plasmid COLE2imm-K317 DNA, replicon region
encoding Rep protein.
ACCESSION D30058
NID 9487266
VERSION D30058.1 GI:487266
KEYWORDS Rep protein; autonomous replication; initiator protein; replicon.
SOURCE Escherichia coli plasmid:COLE2imm-K317 DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 1457)
Itoh,T.
Direct Submission
Submitted (28-APR-1994) to the DDBJ/EMBL/Genbank databases. Taro
Itoh, Osaka University, Faculty of Science, Department of Biology;
1-1 Machikaneyama-cho, Toyonaka, Osaka 560, Japan
(E-Mail:c62528@center.osaka-u.ac.jp, Tel:06-850-5432(ex.4307),
Fax:06-850-5440)
2 (bases 1 to 1457)
Hiraga,S., Sugiyama,T. and Itoh,T.
Comparative analysis of the replicon regions of eleven
COLE-related plasmids
J Bacteriol. 176 (23), 7233-7243 (1994)
JOURNAL 95050306
MEDLINE
FEATURES
source Location/Qualifiers
1..1457
/organism="Escherichia coli"
/plasmid="plasmid COLE2imm-K317"
/db_xref="taxon:562"
492..1427
/codon_start=1
/transl_table=11
/product="Rep protein"
/protein_id="BAA06296.1"
/db_xref="PID:g1006860"
/db_xref="PID:g809495"
/db_xref="GI:809495"
/translation="MSAALQYFENLPHRPYTTDLAFGLRISGRALLARYIOOQ
PAAQFWLVDVDRGGAIDMSDRNAPATVKNFVNGHAHLVALNTAVRPAADASY
KALKYAAVERSLCEKLCADVYSGLICNPHLEMLVMEEREAVYITDELADYDLS
ASARSIDKHVGMGRNCHLEMTKRAYRAIRGMPAFSOWMDAVIOHREVMNASLPV
PLSPPECAIGKSIKTYTRNTPTPTFQVYADTTPPELQAKRGKGIACEALYDVK
RMAALCMLENGISOKRAIAMLLEVSRTTRNMSGR"

BASE COUNT 361 a 359 c 386 g 351 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 1457;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1355 TTCTGAGAA 1363
 Cp 33 ttctcgagaa 25

RESULT 71
 LOCUS EIN9981 1477 bp DNA BCT 07-JAN-1999
 DEFINITION Enterobacter Intermedius plasmid pLV1402 alpha replicon, partial.
 ACCESSION AJ009981
 MID 94127797
 VERSION AJ009981.1 GI:4127797
 KEYWORDS copA gene; copB gene; plasmid; repA gene; tapA gene.
 SOURCE Enterobacter intermedius.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Enterobacter.
 REFERENCE 1 (bases 1 to 1477)
 AUTHORS Osborn, A.M.
 JOURNAL Direct Submission
 TITLE Submitted (31-JUL-1998) Osborn A.M., Division of Microbiology, GBF, Division of Microbiology, GBF, Mascheroder Weg 1, Braunschweig, D-38124, GERMANY
 REFERENCE 2 (bases 1 to 1477)
 AUTHORS Osborn, A.M., da Silva-Tatley, F.M., Steyn, L.M., Pickup, R.W. and Saunders, J.R.
 JOURNAL Mosaic plasmids and mosaic replicons: Evolutionary lessons from the analysis of genetic diversity in IncFII-related replicons unpublished
 TITLE 3 (bases 1 to 1477)
 REFERENCE Osborn, A.M., Pickup, R.W. and Saunders, J.R.
 AUTHORS Plasmid replicons, from lakewater sediments, homologous to the pGSH500 beta replicon are members of a larger group of classa theta replicons widely distributed in bacteria
 JOURNAL Unpublished
 TITLE Location/Qualifiers
 FEATURES
 source
 1. 1477
 /organism="Enterobacter intermedius"
 /plasmid="pLV1402"
 /isolate="AH14"
 /db_xref="taxon:61648"
 1. 200
 /gene="copB"
 <1. 200
 /gene="copB"
 /function="Replication control"
 /codon_start=3
 /transl_table=11
 /product="CopB"
 /protein_id="CA08928.1"
 /db_xref="PID:e1363553"
 /db_xref="PID:g4127798"
 /db_xref="GI:4127798"
 /translation="NPMTASEKQLAAVARKRTKREKVFIRNPLKDMIALCDQEGITQAOPIEKIRERLESEGLK"
 complement(317..>403)
 /gene="copA"
 /function="antisense control"
 complement(317..403)
 /gene="copA"
 421. 498
 /gene="copA"
 421. 498
 /gene="tapA"
 421. 498
 /gene="tapA"
 /function="antisense control"
 /codon_start=1
 /transl_table=11
 /product="TapA"
 /protein_id="CA08929.1"
 /db_xref="PID:e1363554"
 /db_xref="PID:g4127799"
 /db_xref="GI:4127799"

gene
 CDS
 /translation="MLRKIQYLFLLPCNISAGRCD"
 491..1354
 /gene="repA"
 491..1354
 /gene="repA"
 /function="Replication protein"
 /codon_start=1
 /transl_table=11
 /product="RepA"
 /protein_id="CA08930.1"
 /db_xref="PID:e1363555"
 /db_xref="PID:g4127800"
 /db_xref="GI:4127800"
 /translation="MTDAOTLFSHYRVKNPNPVPFPRECKTLTPFCRLMAKAEFTSRDFAVHVARMSSEGRKRKMPPLLRRSIDALOGCFFHDDPLANVORSITNLAIECGLATESKKNLSITRATRLKFMALGLITYETEDPOGCIPTDITTPALFSLDSDVAVAARSRSEVENQRRKONKPLEMBELIKRFRVFRFRSIOSEKRGKLRARARADRRKIDIVTLVKQOLREYASGRFTGLDMKRRELERVVERMLMSKGNNYRLATVPI"
 BASE COUNT 355 a 388 c 404 g 330 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 17; Length 1477;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 TTCTGAGAA 25
 Oy 25 ttctcgagaa 33
 RESULT 72
 LOCUS 114026 1482 bp DNA PAT 30-AUG-1995
 DEFINITION Sequence 5 from patent US 5444153.
 ACCESSION 114026
 MID 9996449
 VERSION 114026.1 GI:9996449
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1482)
 AUTHORS Goss, N.H. and Richardson, M.A.
 TITLE Variants of PAI-2
 JOURNAL Patent: US 5444153-A 5 22-AUG-1995;
 FEATURES
 source
 1. 1482
 /organism="unknown"
 BASE COUNT 439 a 307 c 320 g 416 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 1482;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 890 TTCTGAGAA 898
 Oy 25 ttctcgagaa 33
 RESULT 73
 LOCUS AF114501 1489 bp DNA BCT 12-MAR-1999
 DEFINITION Uncultured gamma proteobacterium SUR-ATT-18 complete sequence.
 ACCESSION AF114501
 MID 94406306
 VERSION AF114501.1 GI:4406306
 KEYWORDS
 SOURCE uncultured gamma proteobacterium SUR-ATT-18.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; environmental samples.
 REFERENCE 1 (bases 1 to 1489)
 AUTHORS Actinas, S.G., Anton, J. and Rodriguez-Valera, F.

TITLE Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA

JOURNAL Appl. Environ. Microbiol. 65 (2), 514-522 (1999)

MEDLINE 99124582

REFERENCE 2 (bases 1 to 1489)

AUTHORS Acinas, S.G.

TITLE Direct Submission

JOURNAL Submitted (23-NOV-1998) Microbiology, University Miguel Hernandez, Cita. Valencia Km 87, Alicante, Alicante 03550, Spain

FEATURES

SOURCE

1. 1489

/organism="uncultured gamma proteobacterium SUR-AT-18"

/isolate="SUR ATP"

/db_xref="taxon:86478"

/clone="18"

/note="uncultured attached marine bacterium recovered from surface"

1. 1489

/product="16S ribosomal RNA"

BASE COUNT 391 a 327 c 447 g 324 t

ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 1489;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 278 TTCTGAGAA 286

|||||

QY 25 ttctgagaa 33

RESULT 74

LOCUS E13732 1544 bp DNA PAT 24-JUN-1998

DEFINITION cDNA encoding rat CC chemokine receptor protein.

ACCESSION E13732

NID 93252500

VERSION E13732.1 GI:3252500

KEYWORDS JP 1997227599-A/1.

SOURCE Rattus rattus.

ORGANISM Rattus rattus.

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 1544)

Honda, S. and Fujisawa, T.

CC CHEMOKINE RECEPTOR PROTEIN, ITS PRODUCTION AND USE

Patent: JP 1997227599-A 1 02-SEP-1997;

TAKEDA CHEM IND LTD

OS Rattus rattus (rat)

PN JP 1997227599-A/1

PD 02-SEP-1997

PF 22-FEB-1996 JP 1996035192

PI HONDA SUSUMU, FUJISAWA TOMOYUKI

PC C07K14/705,A61K38/00,A61K38/00,A61K38/00,A61K38/00,A61K38/00,A61K38/00,C07H21/04,C07K16/28,C12N5/10,C12N15/09,C12P21/02,PC C12P21/08//

PC A61K39/395,(C12N5/10,C12R1:91),(C12P21/02,C12R1:91), (C12P21/08,PC C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

Key Location/Qualifiers

FT source 1. 1544

FT /organism="Rattus rattus"

FT /clone="PCR2-PCR"

FT 5'UTR 1. 32

FT CDS 33. 1100

FT /product="Rat CC chemokine receptor protein"

FT 3'UTR 1101. 1544.

FEATURES

Location/Qualifiers

source 1. 1544

/organism="Rattus rattus"

/db_xref="taxon:10117"

BASE COUNT 415 a 347 c 330 g 452 t

ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1544;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1425 TTCTGAGAA 1433

|||||

QY 25 ttctgagaa 33

RESULT 75

LOCUS CHTDMA 1555 bp DNA BCT 28-SEP-1992

DEFINITION C.trachomatis 16S ribosomal RNA gene.

ACCESSION M59178

NID 9144493

VERSION M59178.1 GI:144493

KEYWORDS 16S ribosomal RNA: ribosomal RNA small subunit.

SOURCE C.trachomatis (strain 434) DNA.

ORGANISM Chlamydia trachomatis

Eubacteria; Planctomycetales; Chlamydiaeae; Chlamydia.

1 (bases 1 to 1555)

Weisburg, W.G., Hatch, T. and Woese, C.R.

JOURNAL Unpublished (1990)

FEATURES

Location/Qualifiers

1. 1555

/organism="Chlamydia trachomatis"

/db_xref="taxon:813"

BASE COUNT 401 a 324 c 472 g 358 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 1555;

Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTGAGAA 13

|||||

QY 25 ttctgagaa 33

RESULT 76

LOCUS AB012882 1561 bp mRNA VRT 09-OCT-1998

DEFINITION Cyprinus carpio mRNA for MyoD, complete cds.

ACCESSION AB012882

NID 93721608

VERSION AB012882.1 GI:3721608

KEYWORDS MyoD.

SOURCE Cyprinus carpio cDNA to mRNA.

ORGANISM Cyprinus carpio

Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae; Cyprinus.

1 (bases 1 to 1561)

Watabe, S.

Direct Submission

Submitted (09-APR-1998) to the DDBJ/EMBL/GenBank databases. Shugo Watabe, The University of Tokyo, Graduate School of Agricultural and Life Sciences, Yayoi 1-1-1, Bunkyo, Tokyo 113-8657, Japan (E-mail:awatabe@nongo.ecc.u-tokyo.ac.jp, Tel:81-3-3812-2111(ex.7522), Fax:81-3-5684-0622)

2 (sites)

Kobiyama, A., Mihel, Y., Hirayama, Y., Kikuchi, K., Suetake, H., Johnston, I.A. and Watabe, S.

Molecular cloning and developmental expression patterns of the MyoD and MEF2 families of muscle transcription factors in the carp Unpublished (1998)

Location/Qualifiers

1. 1561

/organism="Cyprinus carpio"

FEATURES

source

CDS

/db_xref="taxon:7962"
191..1018
/codon_start=1
/product="MyoD"
/protein_id="BAA33565.1"
/db_xref="PID:d1034531"
/db_xref="PID:g3721609"
/db_xref="GI:3721609"
/translation="MELSDIPRIPISADDFDPCFNTNDHFEEDLDPRVHSLIK
PDEHHLEDENVRAPSGHHQGRCLLNCKACKRRTTADRRKATMRERRISVND
AFETLRCTSNPNQRLVEILRNALISYIESLQALLGQENNTVPLEHYSQSDAS
SPRNSGDMDFMGPTCOSRRNRNDSSTYEDTNPAAARNKSSVSSLDCLSSIVE
RISTETPACPVLYSEGHGSCSPQESGVSLETGPAPAPPTCPQOQARPIYQVL"

BASE COUNT 446 a 376 c 337 g 402 t

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 1561;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 561 TTCTCAGAA 569
Oy 25 ttctcagaa 33
|||||

RESULT 77
LOCUS AF048817 1563 bp mRNA VRT 18-MAR-1999
DEFINITION Xenopus laevis clone 4 lamina associated protein 2-beta isoform
ACCESSION AF048817
NID 92947305
VERSION AF048817.1 GI:2947305
KEYWORDS
SOURCE African clawed frog.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae;
Xenopus.
1 (bases 1 to 1563)
Gant,T.M., Harris,C.A. and Wilson,K.L.
Roles of LAP2 proteins in nuclear assembly and DNA replication:
truncated LAP2 beta proteins alter lamina assembly, envelope
formation, nuclear size, and DNA replication efficiency in Xenopus
laevis extracts
J. Cell Biol. 144 (6) (1999) In press
2 (bases 1 to 1563)
Gant,T.M. and Wilson,K.L.
Direct Submission
Submitted (17-FEB-1998) Cell Biology and Anatomy, Johns Hopkins
University School of Medicine, 725 N. Wolfe St., Baltimore, MD
21205, USA

FEATURES

Source location/Qualifiers
1..1563
/organism="Xenopus laevis"
/db_xref="taxon:8335"
/clone="4"
1..1563
/note="LAP2 beta isoform"
/codon_start=1
/product="lamina associated protein 2-beta isoform"
/protein_id="AAC05383.1"
/db_xref="PID:g2947306"
/db_xref="GI:2947306"
/translation="MPEFLDPSVLTREKIKSELVANNVTLPSEGRQDVYVQLIOLH
LTSQNRATPDESDEREATPRGRGPRGRATKTTPRAEEDDDVYELSEAL
KEELKYGKPGPILSNTRKLEORLKLREGELSSAPADSSADKNKQNTDSEH
YSDKEEAKIELTFEKRREPLRKSQTQVNRNRTKTEVSDDVTEAAWVSGPAKSG
PVQIVYKELATRTTRPRKVVADPVEDDIAEVPVISEVVEPASNQILTAEN
EHESRKVNAVOPESIKNAETLSVSESELDRSKRLISEKLNLTREDSGSSO
IRLIDVHSGFTNPDLNTALIEETKHIIEELTPKTKTQIKTKVTPYKQIYE
KTFEERTERDLKEMPEPTEPTPGTSGISCRPRTGAAGRLNATDKINNTYSKI
LSNVSRYPAVEVSEKVPGRSLPVMWILFLILVFCLLVYQAMETNBSMFSKL
LGITENSTEN"

BASE COUNT 542 a 316 c 354 g 351 t

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 1563;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 910 TTCTCAGAA 918
CP 33 ttctcagaa 25
|||||

RESULT 78
LOCUS CRANEUR 1569 bp mRNA VRT 07-DEC-1998
DEFINITION Carassius auratus neurolin mRNA, partial cds.
ACCESSION L25056
NID 9407318
VERSION L25056.1 GI:407318
KEYWORDS neurolin.
SOURCE Carassius auratus adult retina, 14 d after optic nerve transection
cDNA to mRNA.
Carassius auratus
Eukaryota; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae; Cyprinidae;
Carassius.
1 (bases 1 to 1569)
Laessing,U., Giordano,S., Stecher,B., Lotzelsch,F. and
Stuerner,C.A.
Molecular characterization of fish neurolin: a growth-associated
cell surface protein and member of the immunoglobulin superfamily
in the fish retinotectal system with similarities to chick protein
DM-GRASP/SC-1/BEN
Differentiation 56 (1-2), 21-29 (1994)
94299040

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 1569;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1371 TTCTCAGAA 1379
CP 33 ttctcagaa 25
|||||

RESULT 79

LOCUS BOVINPHOS 1573 bp mRNA MAM 11-AUG-1994
 DEFINITION Bovine inositol polyphosphate 1-phosphatase (inositol polyphosphate 1-phosphatase gene) mRNA, complete cds.
 ACCESSION M55916
 NID GI:163222
 VERSION M55916.1
 KEYWORDS inositol polyphosphate 1-phosphatase.
 SOURCE Bovine (calf) brain, cDNA to mRNA.
 ORGANISM Bos taurus
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 1573)
 AUTHORS York, J.D. and Majerus, P.W.
 TITLE Isolation and heterologous expression of a cDNA encoding bovine inositol polyphosphate 1-phosphatase
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87, 9548-9552 (1990)
 MEDLINE 91088555
 FEATURES
 source location/Qualifiers
 1..1573
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /tissue_type="brain"
 186..1388
 /gene="inositol polyphosphate 1-phosphatase"
 186..1388
 /gene="inositol polyphosphate 1-phosphatase"
 /codon_start=1
 /product="inositol polyphosphate 1-phosphatase"
 /protein_id="AA0588.1"
 /db_xref="PID:g163223"
 /db_xref="GI:163223"
 /translation="MSDILQELRVSEKANIARACHQETLFPOLLIEKKESEKNNK
 FAYDFKTLADLVQEVKEMENKFGPKKIFGESENELNDLGEKIIWRLGPTTEE
 TVALLSKVNLNKRKASEALAKVYHODVDFPDALDSVEINIPDIDIGIWDPIIDSTYO
 YIKGSADITPNQGITFPSCLOCVYTLIGYIYIOTVPMAGYINDFPVSQDLHTRMKQ
 CSHWGLSTLGNHSLPVPSTRNSEMOGTOIPSESGCSRESVVISTEKETIGA
 LSHVCGEIRFRAAGAKSLCVIIGLADIIYIFSDTFKWDSCAAHILRAMSGMWD
 LKCELENNPTGLDLPOLYVHVGNAGVDMANKGLIAYREKOLETFELSLLOH
 LAPVATHT"

BASE COUNT 414 a 404 c 389 g 366 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 23; Length 1573;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1107 TTCACGAA 1115
 |||||
 Cp 33 ttctcagaa 25

RESULT 80
 LOCUS OC09355 1578 bp mRNA MAM 06-JAN-1995
 DEFINITION Oryctolagus cuniculus protein phosphatase 2A1 B gamma subunit (skeletal muscle isolate) mRNA, complete cds.
 ACCESSION U09355
 NID GI:1619214
 VERSION U09355.1
 KEYWORDS GI:1619214
 SOURCE European rabbit.
 ORGANISM Oryctolagus cuniculus
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1 (bases 1 to 1578)
 AUTHORS Zolnierowicz, S., Csontos, C., Bondor, J., Verin, A., Mundy, M.C. and DePaoli-Roach, A.A.
 TITLE Diversity in the regulatory B-subunits of protein phosphatase 2A: identification of a novel isoform highly expressed in brain
 JOURNAL Biochemistry 33, 11858-11867 (1994)
 MEDLINE 95001896
 REFERENCE 2 (bases 1 to 1578)
 AUTHORS DePaoli-Roach, A.A.

TITLE Direct Submission
 JOURNAL Submitted (03-MAY-1994) Anna A. DePaoli-Roach, Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, IN 46202, USA
 FEATURES
 source location/Qualifiers
 1..1578
 /organism="Oryctolagus cuniculus"
 /strain="New Zealand White"
 /db_xref="taxon:9986"
 /clone_lib="lambda gt11 cDNA libraries, libraries of A.A. DePaoli-Roach"
 /sex="male"
 /tissue_type="skeletal muscle"
 /dev_stage="adult"
 1..183
 184..1527
 /codon_start=1
 /product="protein phosphatase 2A1 B gamma subunit"
 /protein_id="AA058956.1"
 /db_xref="PID:g1619215"
 /db_xref="GI:1619215"
 /translation="MGEDTDFRKINHSFLRDSYVTEADIIISTVEFNTGELLATGDK
 GGRVTFQREPSKAPSPHSGEYDVYSTQSEPEPDYLKSLIEKIKIKMPQON
 AAHSLSTNDKTIKLMKITERDKRREGINKKEEKLKDLSTVLSLOPVLYKPNMLAV
 EVSPRIETANGHTYHINISVNSDEITMSADLILNLMHLAVTDRSENPYDKRANM
 EDLLETITAFSEPHPHCNLFVYSSKSGSLRLCDMAARALCDKHSKLFEEPEDSRSRF
 FSEIISVSADVSESHSGRYMLTRDYLTAVYMDLMEARPIETIYQVHDLRSKLSLYE
 NDCIFPKFECANSGSDVYMTGAYNNFPMFMRNKKRDVTLASRSKPRVLYPPRR
 YGVGGRRRDDISVDSLDTFKILHTAMHPANIIAATNLYIFQDKVNSDVH"

BASE COUNT 346 a 527 c 472 g 233 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 23; Length 1578;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1012 TTCACGAA 1020
 |||||
 Cp 33 ttctcagaa 25

RESULT 81
 LOCUS AR022629 1614 bp DNA PAT 04-DEC-1998
 DEFINITION Sequence 4 from patent US 5792903.
 ACCESSION AR022629
 NID G3976691
 VERSION AR022629.1
 KEYWORDS GI:3976691
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1614)
 AUTHORS Hirschberg, J., Cunningham, F., Xavier Jr. and Gantt, E.
 TITLE Lycopen cyclase gene
 JOURNAL Patent: US 5792903-A 4 11-AUG-1998.
 FEATURES
 source location/Qualifiers
 1..1614
 /organism="unknown"
 BASE COUNT 447 a 288 c 370 g 509 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 1614;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1224 TTCGAGAA 1232
 |||||
 Oy 25 ttctcagaa 33

RESULT 82
 LOCUS A32139 1646 bp mRNA PAT 12-JUL-1996

DEFINITION H.sapiens cyclin A gene.
ACCESSION A32139
NID 91567307
VERSION A32139.1 GI:1567307
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1646)
TITLE NOVEL HUMAN CYCLIN A COMPOSITIONS AND THEIR PRODUCTION METHOD, A
CORRESPONDING NUCLEOTIDE SEQUENCE, A METHOD AND AGENTS FOR
DETECTING OR DIAGNOSING CELL MULTIPLICATION, AND A METHOD AND
AGENTS FOR INHIBITING CELL MULTIPLICATION
JOURNAL Patent: WO 9112324-A 1 22-AUG-1991;
FEATURES
source Location/Qualifiers
1..1646
/organism="Homo sapiens"
/db_xref="taxon:9606"
108..1403
/gene="cyclin A"
108..1403
/gene="cyclin A"
/codon_start=1
/protein_id="CAA02087.1"
/db_xref="PID:e252763"
/db_xref="PID:g1567308"
/db_xref="GI:1567308"
/translation="MLGNSAPGPAITREALGAASIAADGAPRGPEYOPGKGSAPPTTA
DPGALAVLKSNGPRLAEHQPKTRVAPLADLPVNDHVTVPKMSKOPATIHV
DEAEKAKRKRAESOKIERDALAFNSAISLPGPKPLVPIYPMDSFSEPHYDMS
IVLDEKPVSYNEVDYHEDIHITVLRMEVCKPKPVGMKOPDITNSMALIDMIV
EVGREYKONETLILAVYIDRPLISMSVLRGKIQVGTATAMILASKEETIPREVAE
FVITTDITTKOVLRMHVLVKKVITFDLAPVYNOFTQIFLRQAPNCEYSLAE
LGEISLIDADYLYLVPSVIAGAAHLVLYTGQSWPESLIRKTGYTLESKPLCLMD
LHQYTLKAPQAOOSIRKYNKSKHGVSLNPETLNL"

BASE COUNT 477 a 361 c 379 g 429 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1646;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 976 TTCGAGAA 984
OY 25 ttctgagaa 33

RESULT 83
LOCUS 101971 1652 bp ss-DNA PAT 05-MAR-1993
DEFINITION Sequence 2 from Patent US 4795699.
ACCESSION 101971
NID 9269733
VERSION 101971.1 GI:269733
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1652)
AUTHORS Tabor, S. and Richardson, C.C.
TITLE T7 DNA polymerase
JOURNAL Patent: US 4795699-A 2 03-JAN-1989;
President and Fellows of Harvard College;
Cambridge, MA
FEATURES
source Location/Qualifiers
1..1652
/organism="unknown"
BASE COUNT 435 a 394 c 383 g 440 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1652;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1152 TTCGAGAA 1160
OY 25 ttctgagaa 33

RESULT 84
LOCUS E13663 1658 bp DNA PAT 24-JUN-1998
DEFINITION cDNA encoding cytochrome P450 which is induced by jasmonate.
ACCESSION E13663
NID 93252440
VERSION E13663.1 GI:3252440
KEYWORDS JP 1997224671-A/1.
SOURCE Glycine max.
ORGANISM Glycine max.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphylliphytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Rosidae; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Shibata, D., Kato, T. and Ota, H.
TITLE DNA CODING NEW CYTOCHROME P450
JOURNAL Patent: JP 1997224671-A 1 02-SEP-1997;
MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO:KK
OS Glycine max (soybean)
PN JP 1997224671-A/1
PD 02-SEP-1997
PF 19-FEB-1996 JP 1996031075
PI SHIBATA DAISUKE, KATO TOMOHKO, OTA HIROYUKI
PC C12N15/09, C12N9/02, C12N9/02, C12R1:91;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FH source 1..1658
/organism="Glycine max"
/strain="SB-P"
FT 5'UTR 1..24
FT CDS 25..1551
/product="cytochrome P450"
/gene="CYP93A1"
FT 3'UTR 1552..1658.
FEATURES
source Location/Qualifiers
1..1658
/organism="Glycine max"
/db_xref="taxon:3847"
BASE COUNT 494 a 357 c 393 g 414 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1658;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 600 TTCGAGAA 608
OY 25 ttctgagaa 33

RESULT 85
LOCUS RABUGT2B 1671 bp mRNA MAM 11-MAY-1994
DEFINITION Oryctolagus cuniculus UDP-glucuronosyltransferase (UGT2B14) mRNA,
complete cds.
ACCESSION L01082
NID 9165798
VERSION L01082.1 GI:165798
KEYWORDS 4-hydroxydiphenol; UDP-glucuronosyltransferase.
SOURCE Oryctolagus cuniculus (strain New Zealand White) adult liver mRNA.
ORGANISM Oryctolagus cuniculus
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 1671)

AUTHORS Tukey, R.H., Pendurthi, U.R., Nguyen, N.T., Green, M.D. and Tephly, T.R.
 TITLE Cloning and characterization of rabbit liver
 UDP-glucuronosyltransferase cDNAs. Developmental and Inducible
 Expression of 4-hydroxydiphenyl UGT2B13
 JOURNAL J. Biol. Chem. 268, 15260-15266 (1993)
 MEDLINE 93315511
 FEATURES
 SOURCE
 1. .1671
 /organism="Oryctolagus cuniculus"
 /db_xref="taxon:9986"
 mRNA
 1. .1671
 /gene="UGT2B14"
 1. .28
 /gene="UGT2B14"
 1. .1671
 /gene="UGT2B14"
 29. .1621
 /gene="UGT2B14"
 /codon_start=1
 /product="UDP-glucuronosyltransferase"
 /protein_id="AA18021.1"
 /db_xref="PID:9165799"
 /db_xref="GI:165799"
 /translation="MSYKHYSLVLLLLQLSCCFRTGCGKVLVPMDFSLMNLANTL
 DELVRRGHEVIVLNSASIFIDPSKQNIFFETPIATKDLEDLFVHYSTWJAR
 QNSQMKYFSLQKLFSEYSDSCENACKVEYFNKTLMTKDESDILSDAIGPCEL
 LAELLRIPFYSLRFTPTGMEKYSGLSVSPVPIILSDSGKTFEMRVNMLCM
 LYPDFWFMQNKMKDQYSEVIGRPVTFSELDGKADMTLIRSYWDLFEPRPLPNO
 FVGLHCKPAKPLPKEMEFVOSGEGVYVSLGSVNSMTERANLISAFAPLPO
 KVIWRFQGRPETIGPNTRIYDIPQNDLGHPTKAFVYHGGANGYLKELHIGIPV
 GLPLFGEOPDNIATKGAIRLNMKTSSEDLLNLTIVINDPSTKEVNTLSIH
 HDQPMKPLDRAVEIVYMRKAKHLRVAHDLTWQYHSLDVGFLVSCAFLLEL
 VIKSYLVYQKLVKIKKKORD"
 1639. .1671
 /gene="UGT2B14"
 BASE COUNT 462 a 352 c 384 g 473 t
 ORIGIN
 3' UTR
 Query Match 100.0%; Score 9; DB 23; Length 1671;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 195 TTCGACAA 203
 |||||
 QY 25 ttctcgaa 33
 RESULT 86
 LOCUS PTALPHPT1 1689 bp DNA INV 17-APR-1997
 DEFINITION P. tetraurelia alphaP71 gene.
 ACCESSION X99489
 X1460089
 X99489.1 GI:1460089
 VERSION alpha-tubulin: alphaP71 gene.
 SOURCE Paramaecium tetraurelia.
 ORGANISM Paramaecium tetraurelia
 Eukaryota; Alveolata; Ciliophora; Nassophorea; Periculiada;
 Paramacium.
 1 (bases 1 to 1689)
 Dupuis-Williams, P., Klotz, C., Mazarquill, H. and Belsson, J.
 The tubulin gene family of Paramacium: characterization and
 expression of the alphaP71 and alphaP72 genes which code for
 alpha-tubulins with unusual C-terminal amino acids, GLY and ALA
 Biol. Cell 87, 83-93 (1996)
 2 (bases 1 to 1689)
 Dupuis-Williams, P.
 Direct Submission
 Submitted (22-JUL-1996) P. Dupuis-Williams, Neurobiologie de la
 Diversite Cellulaire, ESPCI, 10 rue Vauquelin, F-75005 PARIS,
 FRANCE
 Related sequence: Science 266:1688-1691 (1994).
 Location/Qualifiers
 1. .1689.

gene
 CDS
 226. .1575
 /gene="alphaP71"
 226. .1575
 /gene="alphaP71"
 /transl_table=6
 /product="alpha-tubulin"
 /protein_id="CAA67847.1"
 /db_xref="PID:e25298"
 /db_xref="PID:91460090"
 /db_xref="GI:1460090"
 /db_xref="SPTREMBL:Q94712"
 /translation="MREVISIHVGQGGIOVGNACMLPCLFHGIQPDGQMPDRTKIG
 GDPAFTPESEFGAKHPRAVFLDEPVIDEVTGYROLFHEQLISCKEDAN
 FARGHTTKEIVDLCDLRIRKLDNCTGLQFLVFSYGGTGLSGLLERLSD
 YGRKSLGFTTIPSPQSTAVVEPNYSILSTHSLHEDVCMNDNEAIVDICRND
 IDRPYTNLNLIAQVISLSLRFDAANDITEFQTNLPVPIHFMCSYAPI
 SAEKAYHEQLSVAEITNSAFEPANMKACDPHGKYMCSMLYRGDVAPKVNAAIAT
 ITRKRIQVDMCPTGPFKYGINQOPVYVGGDLAKVRACMINSSTAIAVEFRID
 HKFDLMYARAAVHVYVGGMEGSEFSEARDELALENDYEVGIEIETADGEEBEG"
 540 a 317 c 344 g 488 t
 BASE COUNT 540 a 317 c 344 g 488 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 21; Length 1689;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 382 TTCGACAA 390
 |||||
 Cp 33 ttctcgaa 25
 RESULT 87
 LOCUS LMNACHRB 1698 bp mRNA INV 30-JUL-1998
 DEFINITION Locusta migratoria mRNA for nACHR beta subunit.
 ACCESSION AJ000393
 X2764512
 X2764512
 AJ000393.1 GI:2764512
 VERSION beta subunit: nACHR gene: nicotinic acetylcholine receptor.
 KEYWORDS migratory locust.
 SOURCE Locusta migratoria
 ORGANISM Locusta migratoria
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Orthoptera; Caelifera; Acrididae; Oedipodinae; Locusta.
 1 (bases 1 to 1698)
 Stetzer, E.
 Direct Submission
 Submitted (14-JUL-1997) Stetzer E., Institute of Physiological
 Chemistry, Johannes-Gutenberg University of Mainz, Duesbergweg 6,
 55099 Mainz, GERMANY
 2 (bases 1 to 1698)
 Hermen, B., Stetzer, E., Thees, R., Heiermann, R., Schratzenholz, A.,
 Ebblinghaus, U., Kretschmer, A., Methfessel, C., Reinhardt, S. and
 Maellicke, A.
 Neuronal nicotinic receptors in the locust Locusta migratoria.
 Cloning and expression
 J. Biol. Chem. 273 (29), 18394-18404 (1998)
 98325051.
 Location/Qualifiers
 1. .1698
 /organism="Locusta migratoria"
 /db_xref="taxon:7004"
 1. .1491
 /gene="nACHR"
 /product="nicotinic acetylcholine receptor"
 1. .1494
 /gene="nACHR"
 <1. .1494
 /gene="nACHR"
 /codon_start=1
 /product="nicotinic acetylcholine receptor, beta subunit".

/protein_id="CA04055.1"
/db_xref="PID:e1227697"
/db_xref="PID:g2764513"
/db_xref="GI:2764513"
/db_xref="SPREMBL:O46135"
/translation="SEDEBLVRLDFRGYNKLRPVONMTOKEVREGLAVOLLINVA
EKNOIMSNWLRVMDYOLQMDADYGGVLRPPDKVMDYVLENNADQNEV
RYKSNVLIPLNGEVLVWPDPATVYSSCITDIYFPDQOTCMKFGSPNDQVSL
YNNKTPDLSGKSGTWDIEVPAYININIGNPTETDITFIIRKLPYVNI
LPTVLSFLCVLYLPLPGEKTYLGLSILSVLELLSKLPSTSLVLIANT
LLEFIMNTVSLVTVLIIMNFRGPRTHRPPIRAVFLYTLPLMLRPKTRK
WMEMPGAPHPGSPDLPRHLPVAAAAGAGHPAKLEAVEISMLHPNRK
VNRKASAELEPPARESSDLSLSPASKATEAVEFIEHLNEDQYIOIRDMKY
VAMVIDRLQILFELVTTAGTIGLMDAPHFEVLDDRIEIRGR"

BASE COUNT 369 a 498 c 472 g 359 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 1698;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1544 TTCGAGAA 1552
|||||
|||

Cp 33 ttctcagaa 25

RESULT 88
LOCUS APLAPSH 1700 bp ss-DNA INV 29-MAR-1996
DEFINITION Aplysia californica Apshtb1 gene, complete cds.
ACCESSION L43557
NID 9899117
VERSION L43557.1 GI:899117
KEYWORDS G-protein coupled; link phospholipase C; serotonin receptor.
SOURCE Aplysia californica DNA.
ORGANISM Aplysia californica
Eukaryote; mitochondrial eukaryotes; Metazoa; Mollusca;
Gastropoda; Opisthobranchia; Anaspidae; Aplysiidae; Aplysia.
1 (bases 1 to 1700)
Li,X.C., Giot,J.F., Kuhl,D., Hen,R. and Kandel,E.R.
Cloning and characterization of two related serotonergic receptors
from the brain and the reproductive system of Aplysia that activate
phospholipase C
J. Neurosci. 15 (11), 7585-7591 (1995)
96066780

JOURNAL MEDLINE
FEATURES
SOURCE location/Qualifiers
1. 1700
/organism="Aplysia californica"
/db_xref="taxon:6500"
252. 1613
/gene="Apshtb1"
252. 1613
/gene="Apshtb1"
/codon_start=1
/protein_id="AAA93101.1"
/db_xref="PID:g899118"
/db_xref="GI:899118"
/translation="MKSLSKSTHDVPHPEHYVAPAYDEOHLFFSHGVLTIGVGS
LITVAVGVNVLAIETFERILSHSKSNFVIGLAVADLALLVMTFALVMDMY
WIREGTCFIMSDVACERATISICVISTDRKQVOKPLHYEERTTRALLITIC
LWICSVLSFVPIFLMELHSEVEIKAFKDKTEKEKALEHNSALNGLDNDK
SNARKVCLFDVHTGYSVYSFICFYVCTLMLTYLRFLTAQTHQVIRSLQMTNP
QLRGOGASSYRNQGTGSKARLTITITGTLACWLPFFIIPAADEHRLPIECFM
VTIWLGYNSVNPILYGTNSKRAAFKRLRCRSYKSVYSGISPVSPARAFSWIR
PSRLDLSSEHPDACDTGRGKNSKGCCANADPTKPDVSVSEIITAGTVPDSDA
FSS"

BASE COUNT 431 a 416 c 352 g 501 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 1700;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1639 TTCGAGAA 1647

|||||
|||

Oy 25 ttctcagaa 33

RESULT 89
LOCUS A00769 1704 bp DNA PAT 01-DEC-1994
DEFINITION A. hypogaea gene for stilbene synthase.
ACCESSION A00769
NID 958313
VERSION A00769.1 GI:58313
KEYWORDS stilbene synthase.
SOURCE peanut.
ORGANISM Arachis hypogaea
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; endicotyledons;
rosidae; Fabales; Fabaceae; Papilionoideae; Arachis.

REFERENCE
AUTHORS Schroeder,G., Schroeder,J., Hahn,R. and Schreier,P.H.
TITLE Stilbene synthase gene
JOURNAL Patent: EP 0309862-A 1 05-Apr-1989;
BAYER AG
SOURCE location/Qualifiers
1. 1704
/organism="Arachis hypogaea"
/db_xref="taxon:3818"
join(16..193,564..1555)
/codon_start=1
/product="stilbensynthase"
/protein_id="CA00091.1"
/db_xref="PID:g58314"
/db_xref="GI:58314"
/db_xref="SWISS-PROT:P20178"
/translation="MVSQGIKRVORABEGATVLAIGTANPNQVOSTVADYFRYT
NCEHMDIKKKRRCRCRQIKRNMHTFEIILKENNNMAYKAPSDAREDMIRRY
PRGKEKATRAIKEMGQPKTIHLICTTSGVALPQVDELVLGLDPSYRIKMY
HOCFAGATVLRALKDLAENKDAKRVLCSENAVTRFGPNETMDSLVGLFADG
AAALIIQSDVPEVENPLFEIIVSTDDLVNBSGALIGLREVLGFYLNKSVPLDLS
ONINGALSKAFDPLGISDYNSIFWIAHIGRALIDQVEGRVNLPEKMKATRVLSNY
GNSSACVFFIMLRKRSLETGLTGESLDWGVLFQFGPLIETVLRSAAI"
16. 193
/number=1
Intron 194..561
/number=1
exon 564..>1704
/number=2

BASE COUNT 533 a 278 c 350 g 543 t

ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 1704;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 955 TTCGAGAA 963
|||||
|||

Oy 25 ttctcagaa 33

RESULT 90
LOCUS LPACT11 1716 bp mRNA INV 13-JUL-1995
DEFINITION L. polyphemus mRNA for actin.
ACCESSION Z38129
NID 9558672
VERSION Z38129.1 GI:558672
KEYWORDS actin.
SOURCE Atlantic horseshoe crab.
ORGANISM Limulus polyphemus
Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata;
Xiphosura; Limulidae; Limulus.
REFERENCE
AUTHORS May,M., Sanders,M., Garcia,C., Sakai,J. and Matsudaira,P.
TITLE Sequence and domain organization of scruin, an actin-crosslinking
protein in the acrosomal process of Limulus sperm
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1716)
 AUTHORS Way, M.
 TITLE Direct Submission
 JOURNAL Submitted (14-OCT-1994) Michael Way Ph.D, Whitehead Institute for Biomedical Research, 9, Cambridge Center, Cambridge, MA, 02142, USA
 REFERENCE 3 (bases 1 to 1716)
 AUTHORS Way, M., Sanders, M., Garcia, C., Sakai, D. and Matsudaira, P.
 TITLE Sequence and domain organization of scrupin, an actin-cross-linking protein in the acrosomal process of Limulus sperm
 JOURNAL J. Cell Biol. 128 (1-2), 51-60 (1995)
 MEDLINE 95122639
 FEATURES
 source Location/Qualifiers
 1. 1716
 /organism="Limulus polyphemus"
 /db_xref="taxon:6850"
 /sex="Male"
 /clone="ACT11 (Actin)"
 /dev_stage="Adult"
 /tissue_type="testes"
 /clone_lib="Limulus polyphemus testes in lambda ZapII"
 1. 55
 56. 1186
 /product="actin-1"
 /protein_id="CAA86289.1"
 /db_xref="PID:9558673"
 /db_xref="GI:558673"
 /db_xref="SWISS-PROT:P41341"
 /translation="MCEEDVAAALVYDNGSGMCKAGPADAPAPVPSIVGRPHQGV
 MNGGKDSYVGDSEAKSKGILTKPIERHGYTTNDDDEKTHHFFVLELVAPBEH
 PVLTAPLPKANKREKMTQIMETETNPAMVAIQAVLSLHSGSTIVLDGGV
 SHATPIYEGVALPHALRLDLAGRLDVLMLKVTIRGVSFTTAAERIVRIKELIC
 VALDLENEMTTAASSSLEKSYELPDGVITIGNRFRCPEMFPQSLGEMACGIC
 ETEFNIMKCDVIRKDLVANTVLSGSGSTMFGIADRMOKEICAPLAPSTMKIKIILAP
 EKKYSVWIGGSIASLSTFQGMWISKQEDSESPSYVHKCF"
 1187. 1716
 3'UTR 501 a 346 c 362 g 507 t
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 9; DB 21; Length 1716;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1596 TTCTGAGAA 1604
 |||||||||
 Qy 25 ttctgagaa 33

RESULT 91
 LOCUS E02726 1758 bp RNA PAT 29-SEP-1997
 DEFINITION DNA encoding human serum albumin.
 E02726
 ACCESSION
 NID 92170954
 VERSION E02726.1 GI:2170954
 KEYWORDS JP 1991022996-A/2.
 SOURCE synthetic construct.
 ORGANISM
 REFERENCE 1 (bases 1 to 1758)
 AUTHORS Maki, N., Watanabe, K. and Suzuki, M.
 TITLE PRODUCTION OF MATURATION PROTEIN USING CHIMERA LEADER PEPTIDE
 JOURNAL Patent: JP 1991022996-A 2 31-JAN-1991;
 TONEN CORP
 PN JP 1991022996-A/2
 PD 31-JAN-1991
 PF 19-JUN-1989 JP 1989154755
 PI MAKI NOBORU, WATANABE KAZUYA, SUZUKI MASANORI PC
 C12P21/02, C07K7/08, C07K13/00, C07K15/12, C12N1/19, C12N15/62, PC
 C12N15/81,
 PC (C12P21/02, C12R1:645), C07K99:00;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;

CC anti-sense: No;
 CC *source: clone-put-HSA-CH;
 FH key Location/Qualifiers
 FH key
 FT mat-peptide 1. 1755
 FT 1756. 1758.
 FEATURES
 source Location/Qualifiers
 1. 1758
 /organism="Homo sapiens"
 /db_xref="taxon:32630"
 BASE COUNT 546 a 348 c 406 g 458 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 1758;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1548 TTCTGAGAA 1556
 |||||||||
 Qy 25 ttctgagaa 33

RESULT 92
 LOCUS E02715 1758 bp RNA PAT 29-SEP-1997
 DEFINITION cDNA sequence coding for human serum albumin A.
 E02715
 ACCESSION
 NID 92170943
 VERSION E02715.1 GI:2170943
 KEYWORDS JP 1991022984-A/2.
 SOURCE Homo sapiens.
 ORGANISM
 REFERENCE 1 (bases 1 to 1758)
 AUTHORS Maki, N., Watanabe, K. and Suzuki, M.
 TITLE PRODUCTION FOR HUMAN SERUM ALBUMIN A
 JOURNAL Patent: JP 1991022984-A 2 31-JAN-1991;
 TONEN CORP
 OS Homo sapiens (Human)
 PN JP 1991022984-A/2
 PD 31-JAN-1991
 PF 19-JUN-1989 JP 1989154758
 PI MAKI NOBORU, WATANABE KAZUYA, SUZUKI MASANORI PC
 C12N15/14, C12N1/19, C12P21/02, (C12N1/19, C12R1:85), (C12P21/02, PC
 C12R1:85);
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue-type-liver;
 CC *source: clone-LundagLL(HSA-II);
 FH key Location/Qualifiers
 FH key
 FT CDS 1. 1728
 FT /product="Human serum albumin".
 FEATURES
 source Location/Qualifiers
 1. 1758
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 546 a 348 c 406 g 458 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 1758;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1548 TTCTGAGAA 1556
 |||||||||
 Qy 25 ttctgagaa 33

RESULT 93

ID E10761 standard: RNA: HUM: 1761 BP.
AC E10761:
SV E10761.1
NI d1109098
DT 07-OCT-1997 (Rel. 52, Created)
DE 08-OCT-1997 (Rel. 52, Last updated, Version 1)
DE Variant cDNA encoding human serum albumin.
KW JP 1996051982-A/3.
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates;
OC Catarrhini; Homidae; Homo.
RN [1]
RA 1-1761
RP Higashida H., Murakami K., Hama Y., Tsukamoto Y., Isoai A., Kumagai H.:
RT "MODIFIED GENE CODING HUMAN SERUM ALBUMIN".
RL Patent number JP 1996051982-A/3, 27-FEB-1996.
RL ASahi GLASS CO LTD.
CC OS Homo sapiens (human)
CC PN JP 1996051982-A/3
CC PD 27-FEB-1996
CC PF 11-AUG-1994 JP 1994209369
CC PI HIGASHIDA HIDEKI, MURAKAMI KIMIO, HAMA YUKO, TSUKAMOTO YOKO,
CC PI ISOAI ATSUSHI, KUMAGAI HIROMICHI
CC PC C12N5/09,C07K19/00//C12N1/19,C12P21/02,(C12N1/19,C12R1:645),
CC PC (C12P21/02,
CC PC C12R1:645);
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC CC Location/Qualifiers
CC FH Key
CC FH 1. 1761
CC FT source
CC FT /organism="Homo sapiens"
CC FT /clone-lib="human liver cDNA library"
CC FT /clone="pRL2bmc"
CC FT mutation
CC FT replace(1134,"T")
CC FT /product="Original human serum albumin"
CC FT misc-feature
CC FT 1129..1134
CC FT /note="EcoRI recognition site which is at
CC FT between second
CC FT domain and third domain"
CC FT CDS
CC FT 1. 1761
CC FT /product="Human serum albumin"
CC FH Key
CC FH Location/Qualifiers
CC FH 1. 1761
CC FT source
CC FT /organism="Homo sapiens"
CC FT /db_xref="taxon:9606"
CC FT Sequence 1761 BP: 548 A; 351 C; 407 G; 455 T; 0 other;
SQ

Query Match 100.0%; Score 9; DB 11; Length 1761;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1551 TTCTGAGAA 1559
|||||||
QY 25 ttctgagaa 33

RESULT 94
LOCUS BOVCPB 1786 bp RNA MAM 28-APR-1993
DEFINITION Bovine cytochrome P-450(11 beta) mRNA, complete cds.
ACCESSION D00361
NTD 9217534
VERSION D00361.1 GI:217534
KEYWORDS cytochrome; cytochrome P-450.
SOURCE Bovine adrenal cortex, cDNA to mRNA, clone pcP-450(11 beta)-3.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 3 to 1786)
AUTHORS Morohashi, K., Yoshioaka, H., Gotoh, O., Okada, Y., Yamamoto, K.,

TITLE
JOURNAL Miyata, T., Sogawa, K., Fujii-Kuriyama, Y. and Omura, T.
MEDLINE Molecular cloning and nucleotide sequence of DNA of mitochondrial
REFERENCE cytochrome P-450(11 beta) of bovine adrenal cortex
AUTHORS J. Biochem. 102 (3), 559-568 (1987)
88115251
2 (bases 5 to 1531)
Chua, S.C., Szabo, P., Vittek, A., Grzeschik, K.H., John, M. and
White, P.C.
Cloning of cDNA encoding steroid 11 beta-hydroxylase (P450c11)
Proc. Natl. Acad. Sci. U.S.A. 84 (20), 7193-7197 (1987)
88041073
3 (bases 1 to 1786)
Fujii-Kuriyama, Y. and Omura, T.
Expression of two kinds of cytochrome P-450(11 beta) mRNA in bovine
adrenal cortex
J. Biochem. 104 (5), 683-686 (1988)
89174494
The two cDNA clones, pcP-450(11 beta)-2 and pcP-450(11 beta)-3 were
isolated from a cDNA library. pcP-450(11 beta)-2 has already been
sequenced in [J. Biochem. 102, 559-568 (1987)]. Chua et al. also
isolated and sequenced similar
cDNA clones, pB-11s in [2]. Nucleotide sequence determination
revealed the substitutions of 14 nucleotides and 3 amino acids
between pcP-450(11 beta)-2 and -3. And 4 nucleotide and 3 amino
acid substitutions were found in the coding region between
pcP-450(11 beta)-3 and pB-11s. Blotting analysis indicated that at
least two kinds of P-450(11 beta) mRNA were expressed in individual
animals and that at least two kinds of P-450(11 beta) genes exist
in the bovine genome.

FEATURES
source
1. 1786
/organism="Bos taurus"
/db_xref="taxon:9913"
8. 1519
/note="cytochrome P-450(11 beta) precursor"
/codon_start=1
/protein_id="BA00268.1"
/db_xref="PID:d100717"
/db_xref="PID:g217535"
/db_xref="GI:217535"
/translation="MALNKAARYKAGRWLSLHARLLGTGAVAPKRVLPFEMPC
PGRNMRRLQIWKROGSENNHLDHOFQIEGPFRFDVGRHAFVLPEDVRLQO
ADSRPQMIPLWYARQARGHRCQVFLNGPQWRLNDVLSLPALQYTP
VDGARDPFSQTLKARVLQVANGSLDLDAPVFTYTLASLVLYGRLGTLQPP
DSLNFHRLAEAMKSTVQMLVFPRLSRMSTNMSTNMEFEAMDYIFGVANAIORIG
ELAGHPWYSGIVAEILMRADMTLDTIKATIDLTAGSVTDAFTPLMTLEARNP
EVOQAVROESLYAEARRISENPPORATTEPLRLAKETRLRYPGITERRVSSDYL
QNYHPIAGTLVKVLLYSIGRNPVAFARESTHPQWLDROGSGSRFPLHAGFVROC
LGRRYAEVEMLLLHHVKNFLVTEDEDIKMYRFLIMSTPLPTEFRAIO"
85
/note="C in [J. Biochem. 104, 683-686 (1988)], [J.
Biochem. 102, 559-568 (1987)]; a in [2]"
96
/note="T in [J. Biochem. 104, 683-686 (1988)]; c in [J.
Biochem. 102, 559-568 (1987)], [2]"
185
/note="G in [J. Biochem. 104, 683-686 (1988)], [2]; a in
[J. Biochem. 102, 559-568 (1987)]"
324
/note="G in [J. Biochem. 104, 683-686 (1988)]; a in [J.
Biochem. 102, 559-568 (1987)], [2]"
331
/note="T in [J. Biochem. 104, 683-686 (1988)]; c in [J.
Biochem. 102, 559-568 (1987)], [2]"
406
/note="T in [J. Biochem. 104, 683-686 (1988)]; c in [J.
Biochem. 102, 559-568 (1987)], [2]"
579
/note="G in [J. Biochem. 104, 683-686 (1988)], [J.
Biochem. 102, 559-568 (1987)]; c in [2]"
580
/note="C in [J. Biochem. 104, 683-686 (1988)], [J.

KEYWORDS JP 1987275695-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1791)
AUTHORS Marutsein, R., Jiyon, F.M. and Paoro, S.
TITLE PRODUCTION OF MATURE TYPE HUMAN SERUM ALBUMIN
JOURNAL Patent: JP 1987275695-A 1 30-NOV-1987;
JINTEIKA
COMMENT OS human
PN JP 1987275695-A/1
PD 30-NOV-1987
PE 20-FEB-1987 JP 1987037683
PR 21-FEB-1986 FR 86 8602379
PI MARUTSEIN RAITA, JIYON FURANSOMA MAYCO, PAORO SARUMIENTOSU PC
C12P21/02.C07K13/00.C12N15/00.(C12P21/02.C12R1:19): CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue-type=liver;
CC *source: clone=PL53;
FH key Location/Qualifiers
FT CDS 1..1761
FT mat_peptide 1..1758 /product='mature type of human albumin' FT
FEATURES Location/Qualifiers
source 1..1791
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 562 a 356 c 406 g 467 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1791;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1551 TTCGAGAA 1559
|||||||
Qy 25 ttctgagaa 33

RESULT 98
LOCUS I00062 1797 bp ss-DNA PAT 05-MAR-1993
DEFINITION Sequence 3 from Patent US 4920213.
ACCESSION I00062
NID 9313991
VERSION I00062.1 GI:313991
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1797)
AUTHORS Dale, B. and Cordell, B.
TITLE Method and compositions useful in preventing equine influenza
JOURNAL Patent: US 4920213-A 3 24-APR-1990;
Biotechnology Research Partners, Ltd.;
Mountain View, CA
COMMENT On Jul 30, 1993 this sequence version replaced g1:285529.
FEATURES Location/Qualifiers
source 1..1797
/organism='unknown'
BASE COUNT 640 a 340 c 413 g 404 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1797;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1293 TTCGAGAA 1301

Cp 33 ttctcagaa 25
|||||||
RESULT 99
LOCUS GGU30520 1814 bp mRNA VRT 28-AUG-1995
DEFINITION Gallus gallus P311 POU (3.1) mRNA, complete cds.
ACCESSION U30520
NID 9963089
VERSION U30520.1 GI:963089
KEYWORDS chicken.
SOURCE Gallus gallus
ORGANISM Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1814)
AUTHORS Studler, J.M., Glowinski, J. and Levi-Strauss, M.
TITLE An abundant mRNA of the embryonic brain persists at a high level in
cerebellum, hippocampus and olfactory bulb during adulthood
Eur. J. Neurosci. 5 (6), 614-623 (1993)
JOURNAL 94084289
MEDLINE 2 (bases 1 to 1814)
REFERENCE Studler, J.M.
AUTHORS Direct Submission
TITLE Submitted (29-JUN-1995) Jeanne-Marie Studler, INSERM U-114, College
JOURNAL de France, 11 Place Marcelin Berthelot, Paris cedex 05, 75231,
France
FEATURES Location/Qualifiers
source 1..1814
/organism='Gallus gallus'
/db_xref='taxon:9031'
/dev_stage='young adult'
/tissue_type='cerebellum'
63..269
/gene='3.1'
63..269
/gene='3.1'
/note='putative'
/codon_start=1
/product='P311 POU'
/protein_id='AA4902.1'
/db_xref='PID:9963090'
/db_xref='GI:963090'
/translation='MIYOPROTIVSOKVPEPTSGDGFLLKGLPISKEVNRKSESV
EGACMAPVNGDGHFTKINLYTF'
1790..1795
1814
polyA_site
BASE COUNT 512 a 349 c 347 g 606 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 1814;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 881 TTCGAGAA 889
|||||||
Qy 25 ttctgagaa 33

RESULT 100
LOCUS A46379 1830 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 1 from Patent WO9523857.
ACCESSION A46379
NID 92300583
VERSION A46379.1 GI:2300583
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1830)

AUTHORS Kery-Williams, S.M. and Gilbert, S.C.
TITLE YEAST STRAINS AND MODIFIED ALBUMINS
JOURNAL Patent: WO 9523857-A 1 08-SEP-1995;
COMMENT DELTA BIOTECHNOLOGY LTD (GB)
Other publication AU 1818395 950918.
FEATURES
Source
1. 1830
/organism="Homo sapiens"
/db_xref="taxon:9606"
73. 1827
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA02890.1"
/db_xref="PID:e306153"
/db_xref="PID:g2300584"
/db_xref="GI:2300584"
/translation="DAKSEVAFHFRFDLGEENFKALVLAFAQYLAOCPEDEHVKLVN
EYEFATCVADEAENCDKSLHLEFGDKCTVATLEIYEGNADCCAKOEPENECF
LOKHDNPNLPILVRPEVDVMTAFHNEETFLKYLEIARHPFYFAELFFAKR
YKAFTFCCOAAKACILPKLDELDEGASAKRILKASLOKGERAFKMAVAR
LSORFKAEAEVSKLYTDITKYHTECHCDLLECADRDRLKTYICENDSSSLK
ECCEKPLEKSHCIAEYENDEMPADLPISLADVEESKDVCKNTAEADVFLGFLEY
ARRHPDYSVLLRLAKTYETTELKCCAAAPHECAKVEDEKPLVEEONLIKNC
ELFEOLGEYKFORALLVRYTKVPOVSTPLVSVNIGKVSCKKHPEAKMPCAE
DYLVSVMOLCVLHEKTPVSDRVTKCTESTLVNRRPCELSLEYDEVVPEFNAEFT
FHADICLSKERQIKKOTALVELVNHKPKATKEQLKAVMDFAAEVCKKADKET
CFABEGKRLVAAQNALGL"
BASE COUNT 562 a 363 c 420 g 485 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1830;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1620 TTCTGAGAA 1628
|||||
OY 25 ttctgagaa 33
RESULT 101
LOCUS 117765 1863 bp DNA PAT 28-FEB-1996
DEFINITION Sequence 2 from patent US 5494806.
ACCESSION 117765
NID 91598120
VERSION 117765.1 GI:1598120
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1863)
AUTHORS Segre, G.V., Kronenberg, H.M., Abou-Samra, A., Juppner, H., Potts, J.T.,
Jr. and Schipani, E.
TITLE DNA and vectors encoding the parathyroid hormone receptor,
transformed cells, and recombinant production of PTHrP proteins and
peptides
JOURNAL Patent: US 5494806-A 2 27-FEB-1996;
FEATURES
Source
1. 1863
/organism="unknown"
BASE COUNT 408 a 493 c 522 g 438 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 1863;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1714 TTCTGAGAA 1722
|||||
OY 25 ttctgagaa 33
RESULT 102

LOCUS AF026812 1863 bp DNA BCT 03-NOV-1997
DEFINITION Butyrivibrio fibrisolvens strain OR7 flagellin B protein (flab)
gene, complete cds.
ACCESSION AF026812
NID 92583087
VERSION AF026812.1 GI:2583087
KEYWORDS
SOURCE Butyrivibrio fibrisolvens.
ORGANISM Butyrivibrio fibrisolvens.
REFERENCE 1 (bases 1 to 1863)
AUTHORS Karmokoff, M.L., Allard, S., Austin, J.W., Whitford, M., Heford, M.A.
and Teacher, R.M.
TITLE Isolation, sequencing, and transcriptional analysis of the
flagellins from the rumen anaerobe Butyrivibrio fibrisolvens OR7:
evidence supporting post-translational modification to these
flagellins
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1863)
AUTHORS Karmokoff, M.L., Allard, S., Austin, J.W., Whitford, M., Heford, M.A.
and Teacher, R.M.
TITLE Direct Submission
SUBMITTED (25-SEP-1997) Center for Food and Animal Research,
Agriculture and Agri-Food Canada, Rm. 2046, K. W. Neelby Building,
Ottawa, ON K1A 0C6, Canada
FEATURES
Source
1. 1863
/organism="Butyrivibrio fibrisolvens"
/strain="OR7"
/db_xref="taxon:831"
-35_signal 119. 122
-10_signal 140. 144
RBS 152. 1863
287. 290
/gene="flab"
287. 1826
/gene="flab"
300. 1826
CDS
/gene="flab"
/note="flagellin B structural protein"
/codon_start=1
/transl_table=11
/product="flab"
/protein_id="AAB82613.1"
/db_xref="PID:g2583088"
/db_xref="GI:2583088"
/translation="MYVOHNNQAASSRTLTITTTDAOSKSTKSSGKYKINRAADAA
GUSISEKMKQIRGLGQASTNASDGVSVQTRREGALTEVHDLQRMNLAIVOANGTN
SESDSDSDIONEIDLOEIDRIASTYKNEFFLKGDLGRMYINAHADAGLDGLIQ
NTRATFTMESLEAGEERYTIGOTYTIASSTNOEYAEALHLQYTTAATPTAATK
LWAGTISIDGKTYTINADIGTIDANKVTNGYLAKLITGSEYVYKNGNEYTRSTS
YNSSGIDDAVNTLTKATAAKVMVAILXKKAASSIGATGPAAMDPAKTTSDGKVKAG
DYIMKHNKTKVDSKEITSTYSSTIKGPNVIONALTINLHVGAADAMTKIVTL
EMAKNSIGIAGINSDATGKNATTAIDATIDATDVRSARQRELAQVNRLEHSKTL
DNVVRNTEAESRIRDTOMADTMVEYSNNNIIQAGQSMLOANQANQATGCVLQLO

```

VERSION      AF084366.1 GI:3450894
KEYWORDS
SOURCE       Lactobacillus bulgaricus.
ORGANISM     Lactobacillus bulgaricus.
REFERENCE    1 (bases 1 to 1979)
AUTHORS      Ishino, Y., Kim, S. I. and Soll, D.
TITLE        Direct Submission
JOURNAL      Submitted (14-AUG-1998) Department of Molecular Biophysics &
              Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue,
              New Haven, CT 06520-8114, USA
FEATURES
  source      Location/Qualifiers
              1..1979
              /organism="Lactobacillus bulgaricus"
              /db_xref="taxon:1594"
              411..1913
              /codon_start=1
              /transl_table=11
              /product="glutamyl-tRNA synthetase"
              /protein_id="AAC32611.1"
              /db_xref="PID:93450895"
              /db_xref="GI:3450895"
              /translation="MANKIRVYRAPSPTGHLIGNARTALENTLEFARNKGTLYLRI
              EDADTERNVEGAEISOENLHWLGIDMEGPDIGDGYPKOSERKDIYQYIDDLLE
              EGKAYSEKTEBELPAOREEORAMGIADHYEYEGMTDEIKQAARAKGLKPV
              RHIDPEGVTEMDQIVKGLHSFSDTIGDPIQKRDMPYTNFAVVIDDHLMEISHV
              LRGDHISNTPROLCVYFALGHEAPYFPHMTLINSAGKSKSRDESVLOFIEYRE
              LVSCQPCSTSSSLGMSVGESEISREFIKQDPARLSKSPAFQOKLDMVNNQ
              YMKTDRELDELALNLOEAGLVANAPGMEVNRQLVNNYANOMSYTKQIDVLSK
              IFFTEAKYLTDEVEVEIKDEARPAIEEFKQDLKIDNTAKIKGAIAMATRRRTGK
              GRKLEPMRIRATIRSMVGPICGEAMELMGKIDVYMHDLTLKOLSEAGIE"
BASE COUNT   545 a 492 c 507 g 435 t
ORIGIN
Query Match  100.0%; Score 9; DB 18; Length 1979;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 532 TTCTGAGAA 540
OY 25 ttctgagaa 33

RESULT 109
LOCUS       150943      1980 bp      DNA
DEFINITION Sequence 12 from patent US 5643770.
ACCESSION   150943
NID         92472646
VERSION     150943.1 GI:2472646
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    Unclassified.
AUTHORS      1 (bases 1 to 1980)
TITLE        Mason, J.M. and Squinto, S.P.
JOURNAL      Retrieval vector particles expressing complement inhibitor
              activity
              Patent: US 5643770-A 12-01-JUL-1997;
              Location/Qualifiers
              1..1980
              /organism="unknown"
BASE COUNT   658 a 326 c 317 g 679 t
ORIGIN
Query Match  100.0%; Score 9; DB 25; Length 1980;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 340 TTCTGAGAA 348
OY 25 ttctgagaa 33

```

```

RESULT 110
LOCUS       NU80061      1980 bp      RNA
DEFINITION Naegleria jamesoni 18S ribosomal RNA gene, partial sequence.
ACCESSION   U80061
NID         92351573
VERSION     92351573
KEYWORDS
SOURCE      Naegleria jamesoni.
ORGANISM    Naegleria jamesoni.
REFERENCE    Eukaryotes; mitochondrial eukaryotes; Heterolobosea;
              Schizopyrenidae; Vahlkampffidae; Naegleria.
AUTHORS      Elnivik, C., Decatur, W.A., Embley, T.M., Vogt, V.M. and Johansen, S.
TITLE        1 (bases 1 to 1980)
JOURNAL      Naegleria nuclear introns contain two group I ribozymes with
              different functions in RNA splicing and processing
              Unpublished
              2 (bases 1 to 1984)
              Embley, T.M. and Dyal, P.L.
              Direct Submission
              Submitted (27-NOV-1996) Zoology, Natural History Museum, Cromwell
              Road, London SW7 5BD, UK
              Location/Qualifiers
              1..1984
              /organism="Naegleria jamesoni"
              /strain="PMPB-6"
              /db_xref="taxon:5764"
              <1..>1984
              /product="18S ribosomal RNA"
BASE COUNT   509 a 416 c 529 g 527 t 3 others
ORIGIN
Query Match  100.0%; Score 9; DB 21; Length 1984;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1223 TTCTGAGAA 1231
Cp 33 ttctgagaa 25

RESULT 111
LOCUS       NAU80057      1984 bp      DNA
DEFINITION Naegleria andersoni 18S ribosomal RNA gene, partial sequence.
ACCESSION   U80057
NID         92351569
VERSION     U80057.1 GI:2351569
KEYWORDS
SOURCE      Naegleria andersoni.
ORGANISM    Naegleria andersoni.
REFERENCE    Eukaryotes; mitochondrial eukaryotes; Heterolobosea;
              Schizopyrenidae; Vahlkampffidae; Naegleria.
AUTHORS      Elnivik, C., Decatur, W.A., Embley, T.M., Vogt, V.M. and Johansen, S.
TITLE        1 (bases 1 to 1984)
JOURNAL      Naegleria nuclear introns contain two group I ribozymes with
              different functions in RNA splicing and processing
              Unpublished
              2 (bases 1 to 1984)
              Embley, T.M. and Dyal, P.L.
              Direct Submission
              Submitted (27-NOV-1996) Zoology, Natural History Museum, Cromwell
              Road, London SW7 5BD, UK
              Location/Qualifiers
              1..1984
              /organism="Naegleria andersoni"
              /strain="PMPB-6"
              /db_xref="taxon:5764"
              <1..>1984
              /product="18S ribosomal RNA"
BASE COUNT   509 a 416 c 529 g 527 t 3 others
ORIGIN
Query Match  100.0%; Score 9; DB 21; Length 1984;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1223 TTCTGAGAA 1231
Cp 33 ttctgagaa 25

```

Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1228 TTCACGAA 1236
|||||||
Cp 33 ttctcagaa 25

RESULT 112
LOCUS ECEPERABCD 1992 bp DNA BCT 07-OCT-1996
DEFINITION E.coli perA, perB, perC and perD genes.
ACCESSION Z48561
NID g1469231
VERSION Z48561.1 GI:1469231
KEYWORDS perA gene; perB gene; perC gene; perD gene.
SOURCE Escherichia coli.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 1990)
AUTHORS Gomez-Duarte,O.G. and Kaper,J.B.
TITLE A plasmid-encoded regulatory region activates chromosomal eaeA
expression in enteropathogenic Escherichia coli
JOURNAL Infect. Immun. 63 (5), 1767-1776 (1995)
MEDLINE 95247259
REFERENCE 2 (bases 1 to 1992)
AUTHORS Kaper,J.B.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-1995) Kaper J. B., Univ. of Maryland School of
Medicine, Center for Vaccine Development, 10 South Pine St.,
Baltimore, MD, USA, 21201
REMARK 3 (bases 1 to 1992)
REFERENCE Kaper,J.B.
AUTHORS Direct Submission
JOURNAL Submitted (28-JUL-1996) Kaper J. B., Univ. of Maryland School of
Medicine, Center for Vaccine Development, 10 South Pine St.,
Baltimore, MD, USA, 21201
COMMENT On Jul 30, 1996 this sequence version replaced g1:695508.
FEATURES
source location/Qualifiers
1..1992
/organism="Escherichia coli"
/strain="enteropathogenic E. coli"
/isolate="E2348/69"
/db_xref="taxon:562"
/clone="pOG020"
56..880
/gene="perA"
56..880
/gene="perA"
/function="transcriptional activator"
/note="first ORF of four ORFs that together increase eaeA
transcription. Belongs to the Arc family"
/citation=[1]
/codon_start=1
/transl_table=1
/evidence=experimental
/product="perA"
/protein_id="CA88445.1"
/db_xref="PID:g256055"
/db_xref="PID:g1469232"
/db_xref="GI:1469232"
/db_xref="GI:1469232"
/db_xref="SWISS-PROT:P43459"
/translation="MTSKKMOSSKOEENLALLTNTYISYONIVIFGNOFKIR
NKRETEYTESNSLFLAKNTHMDMEIVGDNPNPYKTIIDALIKLHSISDSDS
CYVKKRTFTNLMOLNIVSNITTDIKYGNKKIKRIIYLSEPDYNDIVNVIS
ASSKSTVDRIKIVTELDISKWKLGVDSSKSPMSDCLRQLNKENITFKKIMDITK
KHASLFLRTDKNIDEISCLGFNSISYIFKVEKYNTTPKRYNGVSIQGLP"
928..1317
/gene="perB"
928..1317
/gene="perB"
/function="transcriptional activator"

/note="second ORF of four ORFs that together increase eaeA
transcription"
/citation=[1]
/codon_start=1
/transl_table=1
/evidence=experimental
/product="perB"
/protein_id="CA88446.1"
/db_xref="PID:g695510"
/db_xref="GI:695510"
/db_xref="SWISS-PROT:P43474"
/translation="MKNNDREKEVDFDCNVLSPSGKAITPKNNATSEIVLF
IPPKAYHILIKNTKICELFSDHIGEDDIISQNAFESHVNHITALVDVANKK
SYASNVKFLITMEGGDILSESKRAP"
1319..1324
/citation=[1]
1342..1347
/citation=[1]
1367..1636
/gene="perC"
1367..1636
/gene="perC"
/function="transcriptional activator"
/note="third ORF of four ORFs that together increase eaeA
transcription"
/citation=[1]
/codon_start=1
/transl_table=1
/evidence=experimental
/product="perC"
/protein_id="CA88447.1"
/db_xref="PID:g695511"
/db_xref="GI:695511"
/db_xref="SWISS-PROT:P43475"
/translation="MEIRBDKAKYLEEKGFRADRNAEIIVLLSDAERKLAQKR
AFCINKSLRNQADNYSIDIKGVYKAYKMGVLNVEKIFRNKEN"
1749..1985
/note="fourth ORF of four ORFs that together increase eaeA
transcription"
/citation=[1]
/codon_start=1
/transl_table=1
/evidence=experimental
/product="perD"
/protein_id="CA88448.1"
/db_xref="PID:g695512"
/db_xref="GI:695512"
/db_xref="SWISS-PROT:P43476"
/translation="MOTINPIPTERRRLKQKTFHKTKKSLFKILCKLSEVTVQVA
VTKLDKKAHCHSPVLOHRHPFLGRKDCQPHRR"
BASE COUNT 767 a 302 c 352 g 571 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 1992;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1857 TTCACGAA 1865
|||||||
Cp 33 ttctcagaa 25

RESULT 113
LOCUS DRU31079 2006 bp mRNA VRT 15-AUG-1996
DEFINITION Danio rerio 47 kDa heat shock protein (hsp) mRNA, complete cds.
ACCESSION U31079
NID g1098626
VERSION U31079.1 GI:1098626
KEYWORDS zebrafish.
SOURCE Danio rerio
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;

Ostariophysi: Cypriniformes; Cyprinidae; Cyprinidae; Rasbora: Danio.

REFERENCE 1 (bases 1 to 2006)
AUTHORS Pearson,D.S., Kulyk,W.M., Kelly,G.M. and Krone,P.H.
TITLE Cloning and characterization of a cDNA encoding the collagen-binding stress protein hsp47 in zebrafish
JOURNAL DNA Cell Biol. 15 (3), 263-272 (1996)
MEDLINE 96226405
REFERENCE 2 (bases 1 to 2006)
AUTHORS Pearson,D.S., Kulyk,W.M. and Krone,P.H.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1995) Patrick Krone, Anatomy and Cell Biology, University of Saskatchewan, Health Science Building, 107 Wiggins Road, Saskatoon, Saskatchewan S7N 5E5, Canada

FEATURES
SOURCE Location/Qualifiers

1..2006 /organism="Danio rerio"
/db_xref="taxon:7955"
5'UTR 1..178
repeat_region 8..49
gene /rpt_unit="GAA 178..1392
/gene="hsp"
178..1392
/gene="hsp"
/note="member of serpin superfamily"
/codon_start=1
/product="47 kDa heat shock protein"
/protein_id="AAB06193.1"
/db_xref="PID:g1098627"
/db_xref="GI:1098627"

/translation="MAYSSLIACLLAAVAGSGEDKRTSTHATSMADTSANLAFLYLN
VKKGLNIFITSPVVAASLIGVMGSKSPASGVKSLIKADALKDHIHTGSEILT
EVSDFQIRNTWKISNRLYGPSSYFPEDEYTKNKKRYNENHSHINRDRSAINSIN
EMAKITDGLPEITKDVKNATDGMATVNAEFKHWDEKHHKAVDNGFLVTSHTV
SVPMHRTGTYGFYEDENRFLVSIALHKKSSMIFIMPEHVEPLDLLENLIRQL
DTWISKLEBAVAISLPKVSMEASHDLQKHGELGLTEAVDKSRADISNKGKDLT
SNVFASSLEWDEGPNFDPISFGESEKRNPKLFYADHPFLVKNKNTSILFGLR
VPRKDKMRDEL"
misc_feature 1377..1399
/note-"encodes RDEL endoplasmic reticulum retention signal"

BASE COUNT 612 a 418 c 442 g 534 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2006;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 50 TTCGTGAGA 58
|||||||
QY 25 ttctgagaa 33

RESULT 114
LOCUS 104790 2015 bp PAT 14-NOV-1994
DEFINITION Sequence 3 from Patent EP 0206733.
ACCESSION 104790
NID 9591462
VERSION 104790.1 GI:591462

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2015)
AUTHORS Burns,A.L.
TITLE Cloned human serum albumin gene
JOURNAL Patent: EP 0206733-A1 3 30-DEC-1986;
FEATURES Location/Qualifiers
SOURCE 1..2015
/organism="unknown"
BASE COUNT 612 a 406 c 447 g 550 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 2015;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1651 TTCGTGAGA 1659
|||||||
QY 25 ttctgagaa 33

RESULT 115
LOCUS A40251 2037 bp DNA PAT 05-MAR-1997
DEFINITION Sequence 1 from Patent WO9424292.
ACCESSION A40251
NID 92296353
VERSION A40251.1 GI:2296353

KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 2037)
AUTHORS Volland,P., Kleczkowski,L., Olsen,O., Poulsen,P., Okkels,F. and Marcussen,J.
TITLE TRANSGENIC ORGANISM
JOURNAL Patent: WO 9424292-A 1 27-OCT-1994;
DANISCO (DK)
COMMENT Other publication GB 2291878 960207
Other publication AU 6539294 941108
Other publication CA 2160159 941027.
Location/Qualifiers

FEATURES
SOURCE 1..2037
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 496 a 557 c 552 g 432 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 2037;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1007 TTCGTGAGA 1015
|||||||
QY 25 ttctgagaa 33

RESULT 116
LOCUS AF031238 2046 bp mRNA VRT 23-MAR-1998
DEFINITION Gallus gallus photoreceptor outer segment membrane glycoprotein (crd41) mRNA, complete cds.
ACCESSION AF031238
NID 92642233
VERSION AF031238.1 GI:2642233

KEYWORDS
SOURCE chicken.
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 2046)
AUTHORS Eukaryotae; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
TITLE Weng,J., Belecky-Adams,T., Adler,R. and Travis,G.H.
JOURNAL Identification of two rds/peripherin homologs in the chick retina invest. Ophthalmol. Vis. Sci. 39 (2), 440-443 (1998)
MEDLINE 98138492

REFERENCE 2 (bases 1 to 2046)
AUTHORS Weng,J. and Travis,G.H.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1997) Psychiatry, UT Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235-9111, USA

FEATURES
SOURCE 1..2046
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue-type="retina"
1..2046
gene

CDS

/gene="crds1"
272..1336
/gene="crds1"
/note="similar to mammalian rds/peripherin"
/codon_start=1
/product="photoreceptor outer segment membrane glycoprotein"
/protein_id="AAC6274.1"
/db_xref="PID:6264234"
/db_xref="GI:2642234"
/translation="MALIKYENQKRVKLAQCLIMNPFVAGIIVSMGLFKIE
LKRSEVMDNSESHFVFNLSILMGILSCAFNGFAGICVDSLDPAFAKWKPLKYL
ALCFENILFLFALICFLRGLSLESLAQLNRSKFRDTPGPCRMKPLIDMLQ
IEFKCGNNGFMDIEFQIOMSNRYLDPSKREVRIDKSNVGRYLVDGVPFSCNPS
PRCICIOYVNNNAHYSDYQTEELNLMGCRREALLHYSSMSMGAVVILYMLE
MSYVGLRLHLSLESIANDPEDECESEGHILENSLKIDILKSLKIKGFNOVEA
GAGGAGEEGKTPALTIVS"

BASE COUNT 569 a 435 c 472 g 570 t

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2046;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1951 TTCTGAGAA 1959
|||||
25 ttctgagaa 33.

RESULT 117
LOCUS MF003906 2053 bp mRNA MAM 28-JUN-1996
DEFINITION Macrosclides proboscideus aldehyde dehydrogenase I, eta-crystallin
ACCESSION U03906
VERSION 01136439
KEYWORDS 003906.1 GI:1136439
SOURCE short-eared elephant shrew.
ORGANISM Macrosclides proboscideus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Macrosclides; Macrosclididae;
Macrosclides.
1 (bases 1 to 2053)
Graham,C., Hodin,J. and Wistow,G.
A retinaldehyde dehydrogenase as a structural protein in a
mammalian eye lens. Gene recruitment of eta-crystallin
J. Biol. Chem. 271 (26), 15623-15628 (1996)
J. Biol. Chem. 271 (26), 15623-15628 (1996)
66279083
2 (bases 1 to 2053)
Wistow,G.J.
Direct Submission
Submitted (02-DEC-1993) Graeme J. Wistow, Molecular Structure and
Function, NEI, NIH, Bethesda, MD 20892 USA
On Dec 28, 1995 this sequence version replaced gi:433360.
Location/Qualifiers
1..2053
/organism="Macrosclides proboscideus"
/db_xref="taxon:29082"
/tissue-type="lens"
/dev_stage="adult"
39..1544
/EC_number="1.2.1.3"
/codon_start=1
/function="enzyme crystallin"
/evidence="experimental"
/product="aldehyde dehydrogenase I, eta-crystallin"
/protein_id="AAC48588.1"
/db_xref="PID:91136440"
/db_xref="GI:1136440"
/translation="MSSSGMPDLPAFLNLIKIOHTRKLFINNEMHDSVSGKTEPFVFNPA
TEEKICEVADKEDVDAKAREAFOMSPRTMDASRGOLIKLADLIERDL
LATLEISNGIKFASAVIMDLQICVIRYACGAMADKIGRTIPVNGEPEFSTRHEPI
GVCGQIFPNAHPHILLACKIGPALCCGNTYIVPAQOTPLTAHVASLIKAGFPFGV

CDS

VNIVPGYPTAGAAISSHMDYDKVAFTEGTEVGMIOEAAKSNILKRYTLELAKNPC
IVFADADDSAVEFPAHOGVFTNOSQSCIAASKLEVEETIYDEFVQSYERAKYVFGN
PLTEPVNHPQINKAQNKKIMELIESGKEGAKLECGGPGNGKGYFOPTIFSNYTD
DMRTAKKEIFPQNOIMKFKSLDIYKANNYYGIVAGVFTKDLDRKAVTVSSALQNG
TWVNCYIAASQSPAGCFKMSGREGVEGIEHYEVKIVTMKISKNS"

BASE COUNT 616 a 396 c 471 g 570 t

ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 2053;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1526 TTCTGAGAA 1534
|||||
25 ttctgagaa 33

RESULT 118
ID E1452 standard; RNA; PLN; 2053 BP.
AC E1452;
SV E1452.1
NI d1109789
DT 07-OCT-1997 (rel. 52, Created)
DT 07-OCT-1997 (rel. 52, Last updated, Version 1)
DE cDNA encoding plant protein which induces meiosis and spore
DE formation.
KW JP 1996140686-A/5.
OS Arabidopsis thaliana (thale cress)
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae;
OC Capriferales; Brassicaceae; Arabidopsis.
RN [1]
RP 1-2053
RA Asada C., Ooto T., Yamamoto M., Shimoda C., Hirayama T.;
RT "GENE CODING FOR VEGETABLE PROTEIN";
RT Patent number JP 1996140686-A/5, 04-JUN-1996.
RL TOTOYA MOTOR CORP.
RL Arabidopsis thaliana
CC PN JP 1996140686-A/5
CC PD 04-JUN-1996
CC PE 25-NOV-1994 JP 1994315514
CC PI ASADA CHIKA, OOTO TOKU, YAMAMOTO MASAYUKI, SHIMODA CHIKA,
CC PI HIRAYAMA TAKASHI
CC PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N5/10;
CC CC strandedness: Double;
CC CC topology: linear;
CC CC hypothetical: No;
CC CC anti-sense: No;
CC FH key Location/Qualifiers
CC FH source 1..2053
CC FT /organism="Arabidopsis thaliana"
CC FT 5'UTR 1..118
CC FT CDS 119..1876
CC FT /product="plant protein which induces meiosis
CC FT and spore
CC FT formation"
CC FT 3'UTR 1877..2053
CC FH key Location/Qualifiers
CC FH source 1..2053
CC FT /organism="Arabidopsis thaliana"
CC FT /db_xref="taxon:3702"
SQ Sequence 2053 BP; 699 A; 330 C; 484 G; 540 T; 0 other;
Query Match 100.0%; Score 9; DB 11; Length 2053;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1018 TTCTGAGAA 1026
|||||
25 ttctgagaa 33


```

CDS
    /sex="male"
    /tissue-type="testis"
    /dev_stage="mature"
    318..1935
    /note=Method: conceptual translation supplied by author"

polyA_site
    /note="12 A nucleotides"
    2112
BASE COUNT      645 a      393 c      435 g      639 t
ORIGIN
Query Match      100.0%; Score 9; DB 23; Length 2112;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2066 TTCTGACAA 2074
|||||
OY 25 ttctgacaa 33

RESULT 122
LOCUS CEU09415 2113 bp mRNA INV 30-NOV-1995
DEFINITION Caenorhabditis elegans putative splicing factor (Ceprp21) mRNA,
complete cds.
ACCESSION U09415
NID 9498845
VERSION U09415.1 GI:498845
KEYWORDS
SOURCE
ORGANISM Caenorhabditis elegans.
          Caenorhabditis elegans
          Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;
          Secernentea; Rhabdillia; Rhabdillidae; Rhabdillinae; Rhabdillidae;
          Rhabdillidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 2113)
AUTHORS Spikes,D.A., Kramer,J., Bingham,P.M. and Van Doren,K.
TITLE SNAP pre-mRNA splicing regulators are a novel, ancient protein
family sharing a highly conserved sequence motif with the prp21
family of constitutive splicing proteins
Nucleic Acids Res. 22 (21), 4510-4519 (1994)
JOURNAL 95061415
MEDLINE 2 (bases 1 to 2113)
REFERENCE Bingham,P.M.
AUTHORS Direct Submission
TITLE Submitted (04-MAY-1994) Paul M. Bingham, State University of New
JOURNAL York at Stony Brook, Life Science Bldg., Rm 450, Stony Brook, NY
11794-5215, USA
FEATURES
Source Location/Qualifiers
1..2113
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone_lib="Caenorhabditis elegans cDNA library, Waterson
R. Nature Genetics 1:114-123(1992)"
15..1982
/gene="Ceprp21"
15..1982
/gene="Ceprp21"
/gene="Ceprp21"
/standard_name="Caenorhabditis elegans pre-mRNA processing
21"
/note="similar to yeast pre-mRNA splicing factor PRP21,"

```

```

Swiss-Prot Accession Number P32524; similar to D.
melanogaster suppressor-of-white-apricot, Swiss-Prot
Accession Number P12297; ubiqlutin-like domain; putative
splicing factor"

/protein_id="AA82164.1"
/clone_lib="AA82164.1"
/db_xref="PIR:G498846"
/translation="MTAAVNSREDSMNNPESLSGATIGLIPPDITIVDTAKF
AAKNGDFENKIREKAKRNPFLSTIDTPYAYKKVYDSSEGRVAPKVOAVKE
HYKRAFVSAPPRAPESADSDSTINAYVDLRLVALFVANGROFLQMTREARN
YQDFDKPACNFTFTKLVDOYQVLPSTVAVQLODDATNKRRLIEDIYRWSME
KIQGKLEKDEEAERKQAVASIMWHDVVQVDFQGGTSQLPPLTPKQVQARI
LLEARNEMKAAAEQEMDESDSDDEDAVQAPAPATAPLPKQKDVIRYDIP
KRVYQKPRAVENWITISPLTGERIPSDLAHVRYKTVDSQYKEDDRHIGRSTEEP
VVALGADIRNLGNFAERTDIEGVGEGOTMIGKLGEDNSOQGNKLIIMPGETEIR
DMITRAVQKLVLLDDQINETHHOGVADPSKEIGAAOVPHOSTGCVTTIGGITT
IPQMSGSPVPPPMGMPKQALPPIQMPFGGSPKRRRTEDDLIPEDMLKTV
NGAISLVNLPQAPERHMGDSIVQFTQVTAPEMSELKQIQDRTGKPVGKRLMSDGL
FYVDNNSAFVYNLADRTAIVLYQVKEGGRK"
123..251
/gene="Ceprp21"
/note="encodes surp module 1"
414..542
/gene="Ceprp21"
/note="encodes surp module 2"
1743..1970
/gene="Ceprp21"
/note="encodes a ubiqlutin-like domain"
2087..2092
polyA_site
2113
/note="43 A residues"
BASE COUNT      649 a      482 c      489 g      493 t
ORIGIN
Query Match      100.0%; Score 9; DB 21; Length 2113;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 270 TTCTGACAA 278
|||||
CP 33 ttctgacaa 25

RESULT 123
LOCUS GDIABR 2123 bp mRNA VRT 09-APR-1991
DEFINITION Chicken mRNA for lamin B receptor.
ACCESSION Y00832
NID 962933
VERSION Y00832.1 GI:62933
KEYWORDS lamin B receptor.
SOURCE chicken.
ORGANISM Gallus gallus
          Gallus gallus
          Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
          Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2123)
AUTHORS Worman,H.J.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1990) H.J. Worman, THE ROCKEFELLER UNIVERSITY,
BOX 292, 1230 YORK AVENUE, NEW YORK NY 10021, USA
REFERENCE 2 (bases 1 to 2123)
AUTHORS Worman,H.J., Evans,C.D. and Blobel,G.
TITLE The lamin B receptor of the nuclear envelope inner membrane: a
polytopic protein with eight potential transmembrane domains
JOURNAL J. Cell Biol. 111 (4), 1535-1542 (1990)
MEDLINE 91009487
FEATURES
Source Location/Qualifiers
1..2123
/organism="Gallus gallus"
/db_xref="taxon:9031"
/dev_stage="embryo"
/clone_lib="lambda zap"
1..2123
mRNA

```


RESULT 129
LOCUS S76880 2202 bp DNA VRT 26-JUL-1995
DEFINITION Xidx-bHLH function dominant negative regulator [Xenopus laevis,
Genomic, 2202 nt].
ACCESSION S76880
NID 9914114
VERSION S76880.1 GI:914114
KEYWORDS
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
Pipridae; Xenopodidae; Xenopus.
REFERENCE
AUTHORS Wilson,R. and Mohun,T.
TITLE Xidx, a dominant negative regulator of bHLH function in early
Xenopus embryos
Mech. Dev. 49 (3), 211-222 (1995)
JOURNAL 95252161
MEDLINE
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gidsbq 163794] from the original journal article.
This sequence comes from Fig. 1A.
FEATURES
source
1..2202
/organism="Xenopus laevis"
/db_xref="taxon:8335"
302..935
/note="Xidx"
/gene="Xidx"
join(302..604,882..935)
/gene="Xidx"
/note="bHLH function dominant negative regulator; This
sequence comes from Fig. 1A"
CDS
/product="Xidx"
/protein_id="AA834225.1"
/db_xref="PID:9914115"
/db_xref="GI:914115"
/translation="MKRISPVRSKSCYQAVCCLSSEGLSIARGSSKRGKGRMPKML
LYDNCCYSKLKELVPGIPQGSKLSQVEILLQHVLDYFDLIVLGEQQSSILSLQK
SDFSELATQDDTVCCH"
BASE COUNT 617 a 485 c 434 g 666 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 2202;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 891 TTCACGAA 899
|||||||
Cp 33 ttctcagaa 25

RESULT 130
LOCUS A30921 2220 bp DNA PAT 14-AUG-1995
DEFINITION VCAM variant DNA from patent WO9200751.
ACCESSION A30921
NID 91247240
VERSION A30921.1 GI:1247240
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE
1 (bases 1 to 2220)
AUTHORS
TITLE A PHARMACEUTICAL COMPOSITION COMPRISING A CELL ADHESION MOLECULE
JOURNAL Patent: WO 9200751-A 11 23-JAN-1992;
FEATURES
source
1..2220
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 578 a 531 c 442 g 669 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 2220;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 539 TTCACGAA 547
|||||||
Cp 33 ttctcagaa 25

RESULT 131
LOCUS CGLY5B 2223 bp mRNA INV 16-APR-1998
DEFINITION Caenorhabditis elegans UDP-galNAc:polypeptide
N-acetyl-galactosaminyltransferase GLY5b (gly-5) mRNA, complete cds.
ACCESSION AF031836
NID 93047192
VERSION AF031836.1 GI:3047192
KEYWORDS
SOURCE
ORGANISM Caenorhabditis elegans.
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS Hagen,F.K. and Nehrke,K.
TITLE 1 (bases 1 to 2223)
CDNA cloning and expression of a family of
UDP-N-acetyl-Dgalactosamine:Polypeptide
N-acetyl-galactosaminyltransferase sequence homologs from
Caenorhabditis elegans
J. Biol. Chem. 273 (14), 8268-8277 (1998)
MEDLINE 98192620
REFERENCE
2 (bases 1 to 2223)
Hagen,F.K. and Nehrke,K.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-1997) Dental Research, University of Rochester,
601 Elmwood Ave, Rochester, NY 14642, USA
FEATURES
source
1..2223
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/chromosome="III"
/note="sequence obtained using EST probes yk72f6 and
cm3e2"
1..2223
/gene="gly-5"
83..1963
/gene="gly-5"
/EC_number="2.4.1.41"
/note="UDP-galNAc:polypeptide
N-acetyl-galactosaminyltransferase splice variant 5b;
pgcArase"
/product="GLY5b"
/protein_id="AAC13672.1"
/db_xref="PID:93047193"
/db_xref="GI:3047193"
/translation="MTIFFKKALKLVLLVPRVNTCSLFFAATNSDSSQIGSNNDLA
MTAEANFHRKAKQDVIOGFGPIPEPEVENNKKEEEOQGNLAKRFENYDNDP
IYKGDAAAGELGKAVVVDKTLSTEAKYDKGMLNNAFNQYASDMLSVHRLPTN
IDAECTERYKNEMLPRTSVLICFHNEMASVLLFTVSHVLEFPTDHLLEEVLYVDFSD
MDHTRKPLEEYNSQFGKVKILMEKREGILRRLGAAVATGEVLTLYDSHCEKNG
NMEPLDRIKRDPTVPCVIDVIDNTEFEYHRSKAYFTSVGSGFDGLQFNHMSIER
DRNKRPRIDPVRSPTMAGLPSIDKEYEKTKTYPGFDINGEULETSFKTMGCG
TLTIVPCHSHVGHFRKRSPTKMTGVNVLKRSIRLAEYWLDDYKTYTERLINQAGD
FGDISRRKRLREDLGCKSFRKYLNDNIYLPFLVPGESVAAGEVANSVQPARCLDCAVG
RHEKRNRPVGTQCHQGGNOYWLSDGELIRDESCVDYAGSDVMVFPCHGMKGNQEM
RYNHDRGLQHAHSQKCLMTKDGAKLEWVACQYDDPYOHMKFEYNEAKAIEHGAKP
PS"
1556..1659
/gene="gly-5"
/note="GLY5b variant segment"

BASE COUNT 656 a 432 c 511 g 624 t
ORIGIN
Query Match 100.0%; Score 9; DB 21; Length 2223;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 654 TTCTCAGAA 662
|||||
Cp 33 ttctcagaa 25

RESULT 132
LOCUS 178692 2225 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 7 from patent US 5693779.
ACCESSION 178692
NID 93014846
VERSION 178692.1 GI:3014846
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2225)
AUTHORS Moos, M. Jr., Krinks, M. and Wang, S.
TITLE Production and use of anti-GorSg12 morphogenetic protein
JOURNAL Patent: US 5693779-A 7 02-DEC-1997;
FEATURES
source 1. 2225
Location/Qualifiers
/organism="unknown"
BASE COUNT 698 a 417 c 425 g 685 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 2225;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1345 TTCTCAGAA 1353
|||||
Cp 25 ttctcagaa 33

RESULT 133
LOCUS 109216 2233 bp PAT 14-NOV-1994
DEFINITION Sequence 3 from Patent WO 8901041.
ACCESSION 109216
NID 9588076
VERSION 109216.1 GI:588076
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2233)
AUTHORS Caras, I. W.
TITLE NUCLEIC ACID AND METHODS FOR THE SYNTHESIS OF NOVEL FUSION
JOURNAL POLYPEPTIDES WITH A PHOSPHOLIPID ANCHOR DOMAIN
FEATURES
source 1. 2233
Location/Qualifiers
/organism="unknown"
BASE COUNT 695 a 455 c 475 g 608 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 2233;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1001 TTCTCAGAA 1009
|||||
Cp 33 ttctcagaa 25

RESULT 134
LOCUS R1RTSM56A 2260 bp DNA BCT 08-JUN-1993

DEFINITION Rickettsia tsutsugamushi 56 kDa type-specific antigen (tsw56) gene,
complete cds.
ACCESSION M63383
NID 9152512
VERSION M63383.1 GI:152512
KEYWORDS Transmembrane protein.
SOURCE Rickettsia tsutsugamushi (strain Kawasaki) DNA.
ORGANISM Rickettsia tsutsugamushi
Eubacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Rickettsiidae; Rickettsia.
REFERENCE 1 (sites)
AUTHORS Tamura, A., Ikeda, H., Nishimoto, H. and Ohashi, N.
TITLE Diversity of immunodominant 56-kDa type-specific antigen (TSA) of
Rickettsia tsutsugamushi: Sequence and comparative analyses of the
genes encoding TSA homologues from four antigenic variants
JOURNAL J. Biol. Chem. 267, 12728-12735 (1992)
FEATURES
source 1. 2260
Location/Qualifiers
/organism="Rickettsia tsutsugamushi"
/strain="Kawasaki"
/db_xref="taxon:784"
79..84
/gene="tsw56"
79..2126
/gene="tsw56"
102..107
/gene="tsw56"
426..431
/gene="tsw56"
430..455
/gene="tsw56"
558..623
/gene="tsw56"
558..2126
/gene="tsw56"
/codon_start=1
/transl_table=1
/product="56 kDa type-specific antigen"
/protein_id="AA26398.1"
/db_xref="pid:9152513"
/db_xref="gi:152513"
/translation="MKIMILASMSALSTPFSASATLEDEGVLECGPVAKIGVNG
VITGESARLPADVDCKKHLSTTLPFGGTLAAGTIAFGFAELGVMYLRINAE
VELGEGKTSQGANAAIDTGAIRKREFLTPPTIMPISIAEDHDLGVDITLQAAV
GQOQLTVEORAEEDRIAMLKNYAGIDYWPDSQNNARVNPVLNITOGAPNVPRP
ONLNLIDHDMRYLVGVLTALSNANKPSVSVKYLSDKITQIYSDIRQFARIANIEVP
GAPLNSASVEIOITKMOELNDVLEELRESFDGLANAFANOIOLNFOIOAOOQOOO
QOQGVYTNQEAALAAAYRALNGNEIOLYKDIYKIDRHAQIRRAMEXLAQESGDDO
SOVSCNDKQQAVALDESKAGSSKSGKKEVLELDMYAOVKLYADVAVATESFETITG
GGVGYARTYGDIDGTSVKHIGVASGVGVAINVADGCVDDIGYIMHSFKIEDKTS
VNAFIANAGVRYNF"
624..2123
/gene="tsw56"
/product="56 kDa type-specific antigen"
BASE COUNT 738 a 297 c 486 g 739 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 2260;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1127 TTCTCAGAA 1135
|||||
Cp 33 ttctcagaa 25

RESULT 135
LOCUS TGLHBB 2308 bp DNA MAM 21-APR-1997
DEFINITION Tachyglusaus aculeatus beta-globin homolog (HBB) gene, complete
cds.
ACCESSION 123800
NID 9438906

```

misc_feature      9..11
                  /note="pot. transcription start sites"
promoter          19..24
                  /note="pot. -10 region"
promoter          69..74
                  /note="pot. alt. -10 region"
misc_feature      78..91
                  /note="pot. alt. transcription start sites"
RBS              99..102
                  /note="pot. ribosome binding site"
CDS              109..969
                  /note="ORF X (AA 1 - 286)"
                  /transl_start=1
                  /transl_table=1
                  /protein_id="CAA32451.1"
                  /db_xref="PID:942721"
                  /db_xref="GI:42721"
                  /db_xref="SWISS-PROT:P23839"
                  /translation="MIRSMYARREIKGEGSATAWERSVNOHLYETFPPEQFES
LEPVVERIRSRITRGVECTLRYPDVSAGGELLNKKLAKOLYTAANVYKMSDEG
EINPVDIRGCVAAQODIDATAETIALDGTLDFTVARETGOALALIBORL
EGVTAEVKVAHSHPEILQMQRESGREVEDAQVQENRLQDELYLAQRIDVAEED
RLEAHVETIYLKKEAVGRRLDFMGEFNRSSNTASKSINAETVNSATELVLE
OMREQIONIE"
repeat_unit       1004..1012
                  /note="inverted repeat A"
terminator        1004..1035
                  /note="pot. stem-loop structure; pot. transcription
repeat_unit       1025..1035
                  /note="inverted repeat A"
promoter          1096..1100
                  /note="35 region"
promoter          1119..1124
                  /note="-10 region"
misc_feature      1130..1131
                  /note="pot. transcription start sites"
RBS              1181..1184
                  /note="pot. ribosome binding site"
CDS              1191..2015
                  /note="ORF Y (AA 1 - 274)"
                  /codon_start=1
                  /transl_table=1
                  /protein_id="CAA32452.1"
                  /db_xref="PID:942722"
                  /db_xref="GI:42722"
                  /db_xref="SWISS-PROT:P23840"
                  /translation="MNEHQPFEEIKLINANGAQWSARQISKLIGYSEYRHPVLVLT
RAKACENSCHTIDHFEELIDNVKIGSNARALKDILVLSRYACLVVYQNDPAKPV
AAGQYFAIOTRROBLADDEAFKQIREDEKRLFELNELKHNKOLVEAAQOAGVAT
DEAFIENHNGYOGIGLDOKAIHQKGLKKNOKLIDHMSVELANLEFRATQTEKIK
RDGVNSKQANTTTFPDGKRVKORTQDEIGGTMPELPPVYSIKOLENSVATTEKK"

BASE COUNT      687 a      553 c      582 g      551 t
ORIGIN
Query Match      100.0%; Score 9; DB 17; Length 2375;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      1594 TTCTGAGAA 1602
      1111111111
OY      25 ttctgagaa 33

RESULT 139
LOCUS      PIGAPOLPB5      2400 bp      DNA      MAM      23-MAR-1995
DEFINITION Sus scrofa apolipoprotein B (allele lpb-5) gene, exons 11-14.
ACCESSION      M20383
NID      9340694
VERSION      M20383.1 GI:340694
KEYWORDS      apolipoprotein B.
SOURCE      Sus scrofa DNA.
ORGANISM      Sus scrofa

```

```

REFERENCE      1 (bases 1 to 2400)
AUTHORS      Maeda,N., Ebert,D.L., Doers,T.M., Newman,M., Hasler-Rapacz,J.,
              Attie,A.D., Rapacz,J. and Smithles,O.
TITLE      Molecular genetics of the apolipoprotein B gene in relation
              to atherosclerosis
JOURNAL      Gene 70 (2), 213-229 (1988)
MEDLINE      89108006
FEATURES
SOURCE
Location/Qualifiers
1..2400
/organism="Sus scrofa"
/db_xref="taxon:9823"
<1..184
/gene="apob"
/number=10
join(185..302,411..557,1426..1637,1899..2136)
/gene="apob"
185..302
/number=11
join(185..302,411..557,1426..1637,1899..2136)
/partial
/gene="apob"
/codon_start=2
/product="apolipoprotein B"
/protein_id="AA64249.1"
/db_xref="PID:9727172"
/db_xref="GI:727172"
/translation="YHRTNPTGTDLLEIADYLLQIRDNCTGNEDHTYLSRYIGNI
GRMEQULPKLTSVLCIKSTOPLLIQKAAIQASRVELGDOVREVLQTEYDVS
PGEKRLAAYLIMRAPSQSDINKYTOQLPGCKNVOKNFVASHANLIHSESYIOEL
KRIYELAKNSQLPTIMDFPKFSNHFSSISLPSLDPVSTTEGULIDPNMYLKP
ESMLKTLRVFGFAPADLFE"
303..410
/gene="apob"
/number=11
411..557
/gene="apob"
/number=12
558..1425
/gene="apob"
1426..1637
/gene="apob"
/number=13
1638..1898
/gene="apob"
/number=13
1899..2136
/gene="apob"
/number=13
2137..>2400
/gene="apob"
/number=14

BASE COUNT      634 a      577 c      478 g      696 t      15 others
ORIGIN
Query Match      100.0%; Score 9; DB 23; Length 2400;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      1248 TTCTGAGAA 1256
      1111111111
OY      25 ttctgagaa 33

RESULT 140
LOCUS      A16054      2409 bp      mRNA      PAT      23-MAR-1994
DEFINITION molecule with the activity of miniactivin.
ACCESSION      A16054
NID      9312777
VERSION      A16054.1 GI:512777

```


Db 2270 TTCGAGAA 2278
|||||
Cp 33 ttctcgagaa 25

RESULT 143
LOCUS AF077838 2466 bp mRNA VRT 13-AUG-1998
DEFINITION Xenopus laevis lamin B1 mRNA, complete cds.
ACCESSION AF077838
NID 93414962
VERSION AF077838.1 GI:3414962
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis
African clawed frog.

REFERENCE
AUTHORS Aaura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
TITLE 1 (bases 1 to 2466)
JOURNAL Dynamic properties of nuclear lamins in Xenopus cells
AUTHORS Takamori, Y. and Inoue, A.
REFERENCE 2 (bases 1 to 2466)
AUTHORS Takamori, Y. and Inoue, A.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-1998) Department of Biology, Graduate School of Science, Osaka University, Machikaneyama-cho 1-1, Toyonaka, Osaka 560-0043, Japan

FEATURES

Source
1..2466
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/cell_type="oocyte"
191..1945
/note="intermediate filament protein; component of the nuclear lamina; located on the nucleoplasmic side of the inner nuclear membrane"
/codon_start=1
/product="lamin B1"
/protein_id="AAC31543.1"
/db_xref="PIR:G414963"
/db_xref="GI:3414963"

CDS

/translation="MATATPSPGPRSSGRSSMTPLSPTRISRLQEKSDLOELNDRIA
LYIDKVSLSSENSLHVOTEREVRSREVSGIKELYELADTRSLDQTAERAK
LOLELSKISVEHODLQASFSKESLEQSAARFETALANALANOSKESKQ
REVELNGENGQLESALAKOLEELILKVDLENCSQLELDPRKNIEEIKR
TRSHETRIEVDGSGROVEYEHKLFQALTEEDHDSQITLYKELEOTSKENR
LASMNSASVSTRELETESRLRINSITSQSEKSRANHDMDQLEMLAKENR
SRKMLAREKEMADIRQMOQQLNDYQLDLVDKALDMEISATKLLEGEBERKLSP
SPSRVTVSRSSASRAVITKGRKRIDVESESSSVLSASATGADVSIEEDVME
GKYIRLNNSKEDPLGGMELTRTIGASVVFKTSRYLAKGQTVIMADAGVTAE
PSSLINKNNSMGTGEDVKATLKNSGEEVAQRTYTTNIPPELEEBEIRPEETA
KESLYSOOKSGNKACAIM"

BASE COUNT 741 a 506 c 634 g 585 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2466;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1567 TTCGAGAA 1575
|||||
Qy 25 ttctcgagaa 33

RESULT 144
LOCUS SS008351 2494 bp mRNA MAN 21-SEP-1994
DEFINITION Sus scrofa vascular cell adhesion molecule mRNA, complete cds.
ACCESSION J008351
NID 9474382
VERSION U008351.1 GI:474382
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

REFERENCE
AUTHORS Vertebra; Eutheria; Artiodactyla; Suiformes; Suidae; Sus.
TITLE 1 (bases 1 to 2494)
JOURNAL Tsang, Y.T., Haakard, D.O. and Robinson, M.K.
MEDLINE Cloning and expression kinetics of porcine vascular cell adhesion molecule
94271236
JOURNAL Biochem. Biophys. Res. Commun. 201 (2), 805-812 (1994)

REFERENCE
AUTHORS Tsang, Y.T.M.
TITLE Direct Submission
JOURNAL Submitted (05-APR-1994) Yvonne T. M. Tsang, Pediatrics (Leukocyte Biology Section), Baylor College of Medicine, Texas Children's Hospital, Clinical Care Center, Suite 1130, 6621 Fannin, Houston, TX 77030-2399, USA

FEATURES

Source
1..2494
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="pv5"
/clone_1lb="porcine aortic endothelial cell cdna library"
/cell_type="endothelial cell"
/tissue_type="aorta"
98..169
/codon_start=1
/product="vascular cell adhesion molecule"
/protein_id="AAA21542.1"
/db_xref="PIR:G474383"
/db_xref="GI:474383"

s1g_peptide

CDS

/translation="MPNIVIFGASNIMLVFAVSQNVKEIFPEDKMTAQIGDSAS
LTCAPDCESLSFMTQIDSPNGVKNKNGRSTLVNMPVSEFENHSLCTVSCGN
LKEGRGLOVEIYFEPKDEIHMSSLEPVGRPVRCIAPVYPERKLEILLDNISM
VSONFELIDIKRKETRSLEFPTPEEDIGKAVCOATLIIGDPSVKTPKRMQVY
ISPRDPIVSNPSTLSQEGDSMMAMTCSGLPAPOISMSKLLNGQQLISGATITL
IAKMDSGIYVEGVNPNVETNKNKEVLTQVAPRODTISVNSSTIESSVMTCS
SDGPAPKILMSKRLKDNLEPSENITLVTSKEDSGIYVEGINQGINRKEVE
LIIQAPKRDQLTAFPESEVKEEDVLIISCTGCVNVPPTLIILKKAETGTVLAKSDG
AVTIHRARLADACVYECESKEKNEIGLOIRSTTLGVKGRSKNDYFSSLLVLCASSLI
IPAIQVITYFARRANMNGSYSLVDAQSKY"
170..1711
/product="vascular cell adhesion molecule"
/note="extracellular domain"
1586..1648
/note="transmembrane domain"
1649..1711
/note="cytoplasmic tail"
1980..1986

BASE COUNT 739 a 500 c 547 g 708 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 2494;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1168 TTCGAGAA 1176
|||||
Qy 25 ttctcgagaa 33

RESULT 145
LOCUS 108629 2500 bp PAT 14-NOV-1994
DEFINITION Sequence 30 from Patent WO 8800205.
ACCESSION 108629
NID 9588672
VERSION 108629.1 GI:588672
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2500)
AUTHORS Wang, E.A., Wozney, J.M. and Rosen, V.A.
TITLE NOVEL OSTEOINDUCTIVE COMPOSITIONS

JOURNAL Patent: WO 8800205-A 30 14-JAN-1988;
 FEATURES Location/Qualifiers
 source 1..2500 /organism="unknown"
 BASE COUNT 507 a 807 c 710 g 476 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 2500;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2036 TTCTCAGAA 2044
 |||||
 25 ttctgagaa 33

QY 25 ttctgagaa 33

RESULT 146
 LOCUS A20130 2580 bp mRNA PAT 23-AUG-1994
 DEFINITION Hybrid insulin/IGF-I receptor mRNA.
 ACCESSION A20130
 NID G583288
 VERSION A20130.1 GI:583288
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 2580)
 AUTHORS
 TITLE A HYBRID CELLULAR RECEPTOR
 JOURNAL Patent: WO 9117252-A 26 14-NOV-1991;
 FEATURES Location/Qualifiers
 source 1..2580 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 597 a 733 c 699 g 551 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 25; Length 2580;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1767 TTCTCAGAA 1775
 |||||
 33 ttctcagaa 25

CP 33 ttctcagaa 25

RESULT 147
 LOCUS LAC0EP 2599 bp DNA BCT 14-JUL-1993
 DEFINITION Lactococcus lactis peptide-binding protein and endopeptidase (oppa) and (pepo) genes, 3' end and complete cds.
 ACCESSION L04938
 NID G293013
 VERSION L04938.1 GI:293013
 KEYWORDS endopeptidase; oppa gene; pepo gene; peptide-binding protein.
 SOURCE Lactococcus lactis (individual isolate P8-2-47, sub-species cremoris) DNA.
 ORGANISM Lactococcus lactis
 Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
 Streptococcaceae; Lactococcus.
 REFERENCE 1 (bases 1 to 2599)
 AUTHORS Mletrau, L., Tan, P. S. T., Haandrikman, A. J., Leenhouts, K. J., Kok, J., Konings, W. N., and Venema, G.
 TITLE Cloning and sequencing of a gene for lactococcal endopeptidase, an enzyme with similarity to mammalian enkephalinase
 JOURNAL J. Bacteriol. 175, 2087-2096 (1993)
 MEDLINE 93209954
 FEATURES Location/Qualifiers
 source 1..2599 /organism="Lactococcus lactis"
 /isolate="P8-2-47"
 /sub_species="cremoris"
 /db_xref="taxon:1358"
 1..375
 gene

CDS
 <1..375
 /gene="oppa"
 /gene="oppa"
 /note="putative"
 /codon_start=1
 /transl_table=11
 /product="peptide-binding protein"
 /protein_id="AAA25203.1"
 /db_xref="PID:9293014"
 /db_xref="GI:283014"
 /translation="DHMTTPPGANDMDITDGSWSIASPQQDFSAAPYNEGFND
 SEITRDLDNDIDSAKSENPYTKAAFEVKYQEDMNKKAYVPYTFNLSLYPVKKRYVGMT
 LDYGMADYMSIEIGVSAKIAIKR"
 488..493
 /gene="pepo"
 /note="putative"
 488..2383
 /gene="pepo"
 500..2383
 /gene="pepo"
 /note="putative"
 /codon_start=1
 /transl_table=11
 /product="endopeptidase"
 /protein_id="AAA25204.1"
 /db_xref="PID:9293015"
 /db_xref="GI:293015"
 /translation="MTRIDDDLEFATVNAEMLENAEIPADKPRISAFDELVLKNEKNLA
 KDLADLSONLPDNDPELLEAIKFKYKAGDMQAREKADPSAVKNELAKVETITDEDFK
 NMLDLYVHSQLPDPFSFSEVPEPMKDAIHISLGSQPLIIPDTTYINDERPKKELL
 DFWANITSEILKTEPVENAEIASALKFDAFLVPSANTSEMAKVAELIYPISTDSF
 VSKVNLDLKSLINDLVKTEPDKYIYEDREESDSLNEENSLKRAMLTYIARG
 ATEFNEEDRLIGGAYGRFLSNVOEAREQEHQDLDESFSQVIGLFYGRKYGEAG
 RADVRMTYAMIKYQARLSKNEISQTAKEALIKDLAIPIGFPEPKLPEIYSRLK
 TTSGLVETDAIKFDELITAPTEKSESDVDTSMHMPAHMNAVYSPSNTIVPPAI
 LQAPRYSLEQSSQNYGICGVIAHEISHARDNNCAQDCEGNLKNKMLDDDYAFEE
 KQKEMIALFDEVELEGANGRLIVSEINAOQGITLALITAKDEKDYDLAIFSQRA
 KIKMRKAKFEQOMLSMDFAPAKLRANIPPTNLEERYDFIVKEIDKMYRAPENRL
 KIW"
 2432..2472
 terminator 874 a 424 c 464 g 837 t
 BASE COUNT
 ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 2599;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1748 TTCTCAGAA 1756
 |||||
 33 ttctcagaa 25

CP 33 ttctcagaa 25

RESULT 148
 LOCUS RABPRLR 2635 bp mRNA MAM 15-SEP-1989
 DEFINITION Rabbit prolactin receptor 2 mRNA, complete cds.
 ACCESSION J04510
 NID G165669
 VERSION J04510.1 GI:165669
 KEYWORDS prolactin receptor.
 SOURCE Rabbit (14-day pregnant) mammary gland, cDNA to mRNA, clones PR8-211[4].
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Eukaryota; Chordata; Metazoa; Chordata;
 Vertebrata; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1 (bases 1 to 2635)
 AUTHORS Ederly, M., Jolicœur, C., Levi-Meynuel, C., Dusanter-Fourt, I., Petridou, B., Boulton, J. M., Lesueur, L., Kelly, P. A., and Djiane, J.
 TITLE Identification and sequence analysis of a second form of prolactin receptor by molecular cloning of complementary DNA from rabbit mammary gland
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86, 2112-2116 (1989)
 MEDLINE 89184578
 COMMENT Draft entry and computer-readable sequence for [1] kindly provided.

by J. Djiane, 14-MAR-1989.

FEATURES

Source Location/Qualifiers

1..2635

/organism="Oryctolagus cuniculus"

/db_xref="taxon:9986"

477..548

/note="prolactin receptor 2 signal peptide"

477..2327

/note="prolactin receptor 2 precursor"

/codon_start=1

/protein_id="AA31457.1"

/db_xref="PID:9165670"

/db_xref="GI:165670"

/translation="MKENYASMTVFLLLFLNIRLLKQSPGPKPIFKRSPEKETETCWRPRGADGGLPTVITLYTHKEGDTITHECPDYTKGDNSTCYFSKRRHSITITITVNATNQMSVSDPRYVDVTVIVPEPPVNTLVEKHHPDRKPTLMKWLPTLVDVRSGLMTLOEIRLKPKEAAEMETHFAGOOQFELSLYPOQKYLVOVRCKPDHGFMSVSPESIQIPNDFTMKDITWVIFAVLSTICIMYAVALKGYSWVTCIFPPVPEPKIKGFDHLEKSESELLSAFGCDPPTADCEDLVEFLVDSDDOOLMPHAKESHSGPMKPTLDPNDSGRGSCDSPILSECEPPOANSPFTPEVIEDEPKRAVHTWDPOTISLVGKMPILSYNGSKSTPPLQPOHNTNSPYHNIAIMCKIATSLDKIDKDALOSSKTEAAGEEKATKOREVSSHSAEDFTGMLPEKPPISPPDLVVEIHKVNKDGLSLLLKQKENGDOGTGAGTPEKSEKAYKSRVMDNMLVLVDDPGQNVALLFESTKEAPPSPQNOAEKDLSTFSTAPSDCRLOQGLDYLDPAFCFMSHLH"

mat_peptide 549..2324

/note="prolactin receptor 2"

BASE COUNT 717 a 668 c 603 g 647 t

ORIGIN 818 bp upstream of BamHI site.

Query Match 100.0%; Score 9; DB 23; Length 2635;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2207 TTCTCGAAG 2215
|||||
Cp 33 ttctcgagaa 25

RESULT 149

LOCUS E12260 2681 bp DNA PAT 24-JUN-1998

DEFINITION cDNA encoding a human homologue of a mouse novel polypeptide derived from stromal cell.

ACCESSION E12260

NID 93251094

VERSION E12260.1 GI:3251094

KEYWORDS JP 1996301898-A/4.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2681)

Hoshino,Y., Nakano,T., Shibayama,S., Tada,H. and Fukushima,D.
NEW POLYPEPTIDE, ITS PRODUCTION, DNA CODING, THE POLYPEPTIDE, VECTOR
COMPRISING THE DNA, HOST CELL TRANSFORMED WITH THE VECTOR, ANTIBODY
OF THE POLYPEPTIDE, AND PHARMACOLOGICAL COMPOSITION CONTAINING THE
PEPTIDE OR ANTIBODY

Patent: JP 1996301898-A 4 19-NOV-1996;

JOURNAL ONO PHARMACEUT CO LTD

COMMENT OS Homo sapiens

PN JP 1996301898-A/4

PD 19-NOV-1996

PF 28-APR-1995 JP 1995128881

PI HONSHIYO YUU, NAKANO TORU, SHIBAYAMA SHIRO, TADA HIDEAKI, PI
FUKUSHIMA DAIRICHI

PC C07K14/47,A61K38/00,A61K38/00,A61K38/00,A61K39/395,
PC A61K39/395,
PC C07H21/04,C07K16/18,C12N1/21,C12N15/09,C12P21/02, PC
C12P21/08//A61K48/00,
PC C12N1/21,C12R1:19),(C12P21/02,C12R1:19);
CC strandedness: single;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;

FEATURES

Source Location/Qualifiers

1..2681

/organism="Homo sapiens"

/db_xref="taxon:9606"

491 t

BASE COUNT 510 a 830 c 850 g

ORIGIN 3' UTR

Query Match 100.0%; Score 9; DB 25; Length 2681;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2357 TTCTCGAAG 2365
|||||
QY 25 ttctcgagaa 33

RESULT 150

LOCUS S72373 2700 bp mRNA VRT 05-JAN-1995

DEFINITION Tlx-tailless homolog [chickens, embryo, mRNA, 2700 nt].

ACCESSION S72373

NID 9619337

VERSION S72373.1 GI:619337

KEYWORDS

SOURCE chicken embryo.

ORGANISM Gallus gallus

Eukaryota; mitochondria eukaryotes; Metazoa; Chordata; Vertebrata; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

1 (bases 1 to 2700)

Tu,R.T., McKeown,M., Evans,R.M. and Umesono,K.
Relationship between Drosophila gap gene tailless and a vertebrate
nuclear receptor Tlx

Nature 370 (6488), 375-379 (1994)

JOURNAL 94322914

MEDLINE

REMARK Genbank staff at the National Library of Medicine created this entry [NCBI gtdbseq 152292]. from the original journal article.
This sequence comes from Fig. 1a.

FEATURES

Source Location/Qualifiers

1..2700

/organism="Gallus gallus"

/db_xref="taxon:9031"

328..1485

/note="tailless homolog; Drosophila terminal/gap gene tailless homolog"

/gene="Tlx"

328..1485

/gene="Tlx"

/note="tailless protein Tll homolog; Drosophila tailless protein Tll homolog. This sequence comes from Fig. 1a."

/codon_start=1

/product="nuclear receptor Tlx"

/protein_id="AAB31467.1"

/db_xref="PID:9619338"

/db_xref="GI:619338"

/translation="MSKPAGSTSRILDPICVCGDRSSGKHYGVYACDGCSEFFERRSI
RRNRTYVCKSGNOGCPVDKTHRNOCRCRLKCLEVNMNRDAVOHERGPTSTIRKQ

VALYRGKKESSGAPHPAPALPAPAFPAFTAVSOLEPHLELAAGTPEKALVGLA
 OPTPKYPRHEVGTMYLYEATVEESVCSAARLLMTSKMAKSVAPFTLSIDQMLL
 EDARRELFVLCIAQMAIPVDANTLLANGNDTDSKINKITSEIOALCOEVARR
 QLRDATEFFACLCIVTKAVPTSGSELKSRNMAAIALADEQOLNYSITRRP
 TOPRFGLLLLRISPSITIEEFKTKIGNVITRLLSDMYKSSDI"

BASE COUNT 723 a 681 c 638 g 658 t
 ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2700;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1089 TTCGAGAA 1097
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 151
 LOCUS AF016539 2704 bp DNA MAM 25-FEB-1998
 DEFINITION Bos taurus enteric beta-defensin (EBD) gene, complete cds.
 ACCESSION AF016539
 NID 92367666
 VERSION AF016539.1 GI:2367666
 KEYWORDS
 SOURCE Bos taurus.
 ORGANISM Bos taurus.
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 2704)
 Gallagher,D.S., Jr., Ryan,A.M., Diamond,G., Bevins,C.L. and
 Womack,J.E.
 Somatic cell mapping of beta-defensin genes to cattle syntenic
 group U25 and fluorescence in situ localization to chromosome 27
 Mamm. Genome 6 (8), 554-556 (1995)
 96014237
 2 (bases 1 to 2704)
 Tarver,A.P., Clark,D.P., Diamond,G., Russell,J.P.,
 Erdjument-Bromage,H., Tempst,P., Cohen,K.S., Jones,D.E.,
 Sweeney,R.W., Mines,M., Hwang,S. and Bevins,C.L.
 Enteric beta-defensin: molecular cloning and characterization of a
 gene with inducible intestinal epithelial cell expression
 associated with Cryptosporidium parvum infection
 Infect. Immun. 66 (3), 1045-1056 (1998)
 98147718
 Erratum: [[published erratum appears in Infect Immun 1998
 May;66(5):2399]]
 3 (bases 1 to 2704)
 Bevins,C.L., Tarver,A.P., Diamond,G., Mines,M. and Hwang,S.
 Direct Submission
 Submitted (30-JUL-1997) Immunology, The Cleveland Clinic
 Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA

FEATURES
 source
 1..2704
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /chromosome="27"
 775..780
 /gene="EBD"
 775..2652
 /gene="EBD"
 join(812..943,2416..2651)
 /gene="EBD"
 /product="enteric beta-defensin"
 812..885
 /gene="EBD"
 812..943
 /gene="EBD"
 827..831
 /gene="EBD"
 859..885
 /gene="EBD"
 join(859..943,2416..2651)
 /gene="EBD"
 5'UTR
 exon
 TATA_signal
 5'UTR
 mRNA
 join(859..943,2416..2651)
 /gene="EBD"

/product="enteric beta-defensin"
 859..943
 /gene="EBD"
 join(886..943,2416..2552)
 /gene="EBD"
 /note="pre-proantimicrobial peptide"
 /codon_start=1
 /product="enteric beta-defensin"
 /protein_id="AAC48804.1"
 /db_xref="PID:92367667"
 /translation="MRLHLLTLFLVLSAGSGFTQGISNPLSCRLNRGICVPICP
 GNLROIGTCFTPSYKCCRMW"
 944..2415
 /gene="EBD"
 2416..2651
 /gene="EBD"
 2553..2651
 /gene="EBD"
 2632..2637
 /gene="EBD"
 2652
 polyA_site
 /gene="EBD"
 682 a 649 c 684 g 689 t
 BASE COUNT 682 a 649 c 684 g 689 t
 ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 2704;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1058 TTCGAGAA 1066
 |||||||||
 Oy 25 ttctcagaa 33

RESULT 152
 LOCUS I13760 2713 bp DNA PAT 16-AUG-1995
 DEFINITION Sequence 1 from patent US 5441881.
 ACCESSION I13760
 NID 9996190
 VERSION I13760.1 GI:996190
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 1 (bases 1 to 2713)
 Deno,T., Ito,H., Kotani,H. and Nakajima,K.
 TITLE NSP7524V restriction-modification genes
 JOURNAL Patent: US 5441881-A 1 15-AUG-1995;
 Location/Qualifiers
 1..2713
 /organism="unknown"
 BASE COUNT 852 a 516 c 427 g 918 t
 ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1352 TTCGAGAA 1360
 |||||||||
 Oy 25 ttctcagaa 33

RESULT 153
 LOCUS CU007694 2713 bp mRNA VRT 11-AUG-1994
 DEFINITION Columba livia prolactin receptor mRNA, complete cds.
 ACCESSION U07694
 NID 9466381
 VERSION U07694.1 GI:466381
 KEYWORDS
 SOURCE domestic pigeon.
 ORGANISM Columba livia

Eukaryotes: mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Columbiformes;
Columbidae; Columba.

REFERENCE 1 (bases 1 to 2713)
AUTHORS Chen, X. and Horseman, N.D.
TITLE Cloning, expression, and mutational analysis of the pigeon prolactin receptor
JOURNAL Endocrinology 135 (1), 269-276 (1994)
MEDLINE 94283267
REFERENCE 2 (bases 1 to 2713)
AUTHORS Chen, X.

Direct Submission
Submitted (14-MAR-1994) Xiaojuan Chen, Physiology & Biophysics,
University of Cincinnati Medical College, 231 Bethesda Ave.,
Cincinnati, OH 45267, USA
FEATURES Location/Qualifiers
SOURCE 1..2713
/organism="Columba livia"
/db_xref="taxon:8932"
/clone="pCLR"
/clone_lib="assembled from a clone of pigeon crop lambda
gt11 cDNA library and a PCR product."
/tissue="crop"
/dev_stage="adult"
1..178
179..2671
/codon_start=1
/product="prolactin receptor"
/protein_id="AA020646.1"
/db_xref="PID:9466382"
/db_xref="GI:466382"
/translation="MKOKLRSSVOIILFALTAVGLTGOSYGRKRIIRCSLEKEIF
SCWKPQSGDGLPNVTLFSSKDSPEKTYPCDPMGPNCSYCPKNTNMTYNT
VMAWETISNSDROYDVITISVOPDAVNTSLERISASTTYLLAKSPPLADYTS
NSHVRYELRLKPEKEWETVSVOQIVNRLQAGVKTVOVRCVLDIGSEMS
SERIHIPNGESPPEKPTIKCRSEKETFCWMPGSDGHPWYLLYKEGERV
YECDPYKAGNSCYFDKHTSFWTIYNIYKATIEIGSNVSDPLYDVITYVQDP
VNTLELKYNRKPYLVLTWSPPLADVRSGMLDYLRLKPEAEWETIFVGOQ
THYMFSLNPKKTYIVQIHCKPDHSGSEMSLEKYLQIPDERIKDMVNIYGVLS
SLICYMTNVIKGYRMIAFILPPVCPKIKGIDTHTLFTGKSELSALGCHGFP
TSDCEELILEYLEDSEDOQLMPSHDHSHSKNKAIAKETSDSGSCDSBSLIS
EKCRSRAILSTLQTDIRDOENNGRRHWETQIASBOKILLENESTKSPWPAQ
LPDNPMPFAHSTVDYHKLCTIDVNIAPVLENEBQHOPOYPIETVLDHNEKR
EVENLYSKTDQTVQKONRPNDSKSPFKLMDEVYEVKRODEVAVALDKHENG
KIEKYTVGTGSKETKSTVDNHLIVLMPDSRIQHIPIVSDPEAMENQNLQGGTEK
NMSYCLVTPSECKETSSEYMDPSFIPARK"

5'UTR
CDS

sig_peptide 179..250
mat_peptide 251..2668
/product="prolactin receptor"
3'UTR 2672..2713
BASE COUNT 907 a 554 c 575 g 677 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2713;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1873 TTCTGAGAA 1881
Oy 25 ttctgagaa 33

RESULT 154
LOCUS CHEKGA 2723 bp DNA VRT 19-MAR-1993
DEFINITION Gallus gallus EGF/EGF-receptor (c-erbB) gene, exons 11-13.
ACCESSION M77638
NID 9211739
VERSION M77638.1 GI:211739
KEYWORDS elongation growth factor receptor; transforming growth factor
receptor.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 2723)
AUTHORS Flickinger, T.W., Mahle, N.J., and Kung, H.-J.
TITLE An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, truncated form of the receptor that can block ligand-dependent transformation
JOURNAL Mol. Cell. Biol. 12, 883-893 (1992)
MEDLINE 92123214
REFERENCE 2 (bases 1 to 2742)

FEATURES Location/Qualifiers
SOURCE 1..2723
/organism="Gallus gallus"
/db_xref="taxon:9031"
1..2723
/gene="c-erbB"
1..301
/partial
/gene="c-erbB"
/note="partial"
/number=10
302..392
/gene="c-erbB"
/number=11
393..852
/gene="c-erbB"
/number=11
853..1052
/gene="c-erbB"
/number=12
1053..1629
/gene="c-erbB"
/number=12
1630..1761
/gene="c-erbB"
/number=13
/note="partial"

exon
intron
exon
intron
exon
intron
BASE COUNT 748 a 567 c 604 g 804 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2723;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 166 TTCTGAGAA 174
Cp 33 ttctcagaa 25

RESULT 155
LOCUS BOVZPF36G 2742 bp DNA MAM 08-AUG-1996
DEFINITION Bos taurus (clone Sal3.8) tristetraprolin (zpf-36) gene, complete
cds.
ACCESSION L42319
NID 91100070
VERSION L42319.1 GI:1100070
KEYWORDS mitogen-inducible gene; tristetraprolin; zinc finger protein.
SOURCE Bos taurus (clone: Sal3.8) DNA.
ORGANISM Bos taurus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 2742)
AUTHORS Taylor, G.A., Thompson, M.J., Lai, M.S. and Blackshear, P.J.
TITLE Phosphorylation of tristetraprolin, a potential zinc finger transcription factor, by mitogen stimulation in intact cells and by mitogen-activated protein kinase in vitro
JOURNAL J. Biol. Chem. 270 (22), 13341-13347 (1995)
MEDLINE 95286626
REFERENCE 2 (bases 1 to 2742)


```

source      1. .2791
            /organism="Bos taurus"
            /db_xref="taxon:9913"
gene        49. .2115
            /gene="beta-ADK2"
            49. .2115
            /gene="beta-ADK2"
            /codon_start=1
            /product="beta-adrenergic receptor kinase 2"
            /protein_id="AA30406.1"
            /db_xref="PID:9162735"
            /db_xref="GI:162735"
            /translation="MADLEAVLADVSYLAMEKSKATPARAKSKIVLPEPSIRSYMO
            KYLEBHEITDKIFNORIGFLLEKDFCLNINEAVPOVKFEYEKEKLENEEDRL
            CSROIVDTYIMKELSCSHPSKQAVHVGSHSKQVSTLPOPYEIEICSLRGS
            IFQKFMESDKTRRCQKKNVLELNLHUTANDSVHRIIRGSGFGEYGRKADTGKMYA
            MKCDKRIRIKKQGETLALNERIMSLVSTGDCPTIVCMITAFHPDKLFTLIDMANG
            GDLYHLSQHGVESEKEMRFATLILGLEMHNFVYVYRLKPANILLDEGHVIRIS
            DLGIACDFSKKRPASVGTGMYAEVYKGTAYSSADWFSGLGCMLEKLLRGSFPR
            OHTKDKHEIDRLMTLNNVELDYVSPSLKLEGLDQVSKRIGSGSAQELKTH
            DFFRGIQWQHYLVQKYPPLIPRGEVNAADAPDGSGFDEEDTGKIKLDDOELYN
            FPLVLSERMOEVAETVYEAVNADTDKTEAKRANKKOLGHEEDYALGRDITVGYML
            KLGNPFLTWQRRFFYLFNRLERGBESQSLTMEQIVSESTQIKDKKILRLI
            KGGQFVYLQCESDEFEVQMKELTETFEAQRLLRRAPKFLNKSRAVVELSKPLCH
            RNSNGL"
BASE COUNT 723 a 692 c 742 g 634 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 2791;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 915 TTCTGAGAA 923
| | | | | | | |
Oy 25 ttctgagaa 33

RESULT 158
LOCUS AB000172 2810 bp mRNA MAM 05-FEB-1999
DEFINITION Porcine mRNA for endopeptidase 24.16, complete cds.
ACCESSION AB000172
NID 91783126
VERSION AB000172.1 GI:1783126
KEYWORDS endopeptidase 24.16 type M2; endopeptidase 24.16 type M3.
SOURCE Sus scrofa Adult Liver cDNA to mRNA.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Suidae; Sus.
REFERENCE
AUTHORS Hirose,S.
TITLE 1 (bases 1 to 2810)
JOURNAL Submitted (27-DEC-1996) to the DDBJ/EMBL/GenBank databases.
Shigenisa Hirose, Tokyo Institute of Technology, Department of
Biological Sciences, 4259 Nagatsuta-cho, Midori-ku, Yokohama,
Kanagawa 226-8501, Japan (E-mail:shirosebio.titech.ac.jp,
Tel:045-924-5726, Fax:045-924-5824)
2 (sites)
REFERENCE Kato,A., Sugitara,N., Saruta,Y., Hosofuji,T., Yasue,H. and Hirose,S.
AUTHORS Targeting of endopeptidase 24.16 to different subcellular
TITLE compartments by alternative promoter usage
JOURNAL J. Biol. Chem. 272 (24), 15313-15322 (1997)
MEDLINE 97326108
FEATURES
source location/Qualifiers
1. .2810
/organism="Sus scrofa"
/db_xref="taxon:9623"
/dev_stage="Adult"
/tissue_type="Liver"
1. .2810
/note="corresponding to exon2,5-16 of this gene;
endopeptidase 24.16 type 2"
130. .2367
/standard_name="endopeptidase 24.16"

```

```

/note="oligopeptidase M :neurolysin :SBAP(soluble
angiotensin-binding protein) :MEP(microsomal
metalloendopeptidase)"
/codon_start=1
/product="endopeptidase 24.16 type M2"
/protein_id="BAA19063.1"
/db_xref="PID:d1019806"
/db_xref="PID:91783127"
/db_xref="GI:1783127"
/translation="MYVPEGLARELATPSSSNAPLGGHPPFVWDCLSCKQGMWSA
RPTNAERSGVGGGILRLRRLTGERMSPLOQNSSTVVDGRNLRDLSPQVSKRRT
ELLTAQTKQVYDDIGMLDIEVYENCLQALADVEVYIVERTMLDPQVSSDKREGRASTA
AASLEAQRSLRSRFDIEMSMREDIFLRIVRLKETCDLQIKPEARRYLEKSKYMKRNG
LHLEAQVQNEIKRAKMKMSSELIDFNKNLNEDDPTFLVSKAEALPDDFIDSLKTD
DNKYITLTKPBYHPYPPVKKCCIPETRRKMEAFNRCKEENTIILOELPLRAVATL
LGYSTHADPLHYEMTAKSTHRTAFDLSQKLKPLCEAEFEFIIKKKECEKPE
YDGINAMDHYTQTEBELKSYDQELKRYFPIVYTBELNITVGBLGSFEQYT
DAHWNNSVTLTYTKDRAVTEVGLQFTLDLPREGKYNHACFGLQGLPLDGRM
VAALVYVNESOPRAGRSLRHDEVRTYFHFEGVMHQICQOTDFARFSGTNTVDY
EVPQMLENNWMDPDSRLRSKHYKDCSPITDILLEKLVASRLVNTGLTLRQIVLSK
VDQSLHNTSLDASAEYAKCTEILGVAAPGTNMPAPFGHLAGYGGYGYLMSYV
FSMDMYSCFRKEGINPVEYMKRNLILKPGSLDGMQLQNLKREPNKATLMR
GLHAP
322. .2367
/standard_name="endopeptidase 24.16"
/note="alternative product for this mRNA: oligopeptidase M
:neurolysin :SBAP(soluble angiotensin-binding protein)
:MEP(microsomal metalloendopeptidase)"
/codon_start=1
/product="endopeptidase 24.16 type M3"
/protein_id="BAA19064.1"
/db_xref="PID:d1019807"
/db_xref="PID:91783128"
/db_xref="GI:1783128"
/translation="MTIGREAMSPLOQNSSTVYGRNVLRDLSPEQIKRTEELIQ
TKQYVDDIGMLDIEVYENCLQALADVEVYIVERTMLDPQVSSDKREGRASTA
DKRLSRREDIEMSMREDIFLRIVRLKETCDLQIKPEARRYLEKSKYMKRNGLPLQ
VQNEIKAMKMKMSSELIDFNKNLNEDDPTFLVSKAEALPDDFIDSLKTDNRKTI
TLKYPKPEVPAKKCCIPETRRKMEAFNRCKEENTIILOELPLRAVATLGYSTH
ADPYLEMNTAKSTHRTAFDLSQKLKPLCEAEFEFIIKKKECEKPEYGGKIN
AMDHYTQTEBELKSYDQELKRYFPIVYTBELNITVGBLGSFEQYTDAHWNN
KSVTLTYTKDRAVTEVGLQFTLDLPREGKYNHACFGLQGLPLDGRMVAALV
VNESOPRAGRSLRHDEVRTYFHFEGVMHQICQOTDFARFSGTNTVDYVPSQM
LENNWMDPDSRLRSKHYKDCSPITDILLEKLVASRLVNTGLTLRQIVLSKYSIH
TNTSLDASAEYAKCTEILGVAAPGTNMPAPFGHLAGYGGYGYLMSYVFSMDMY
YSCFRKEGINPVEYMKRNLILKPGSLDGMQLQNLKREPNKATLMRGLHAP"
BASE COUNT 852 a 558 c 652 g 748 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 2810;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 680 TTCTGAGAA 688
| | | | | | | |
Oy 25 ttctgagaa 33

RESULT 159
LOCUS D88363 2863 bp mRNA VRT 07-FEB-1999
DEFINITION Gallus gallus mRNA for Mltf, complete cds.
ACCESSION D88363
NID 93061279
VERSION D88363.1 GI:3061279
KEYWORDS cm19; Mltf.
SOURCE Gallus gallus (strain:White Leghorn) 9-day-old embryo retinal
pigmented epithelium cDNA to mRNA.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE
AUTHORS Wochil,M.
TITLE Direct Submission

```

JOURNAL Submitted (09-OCT-1996) to the DDBJ/EMBL/Genbank databases. Makoto Mochii, National Institute for Basic Biology, Department of Developmental Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444, Japan (E-mail: mochii@nibb.ac.jp, Tel: +81-564-55-7572, Fax: +81-564-55-7571)

REFERENCE 2 (sites)
Mochii, M., Mazaki, Y., Mizuno, N., Hayashi, H. and Eguchi, G.
Role of Mltf in differentiation and transdifferentiation of chicken pigmented epithelial cell
Dev. Biol. 193 (1), 47-62 (1998)

JOURNAL MEDLINE 98133990

FEATURES
source
location/Qualifiers
1. .2863
/organism="Gallus gallus"
/strain="White Leghorn"
/db_xref="taxon:9031"
/dev_stage="9-day-old embryo"
/tissue_type="retinal pigmented epithelium"
135. .1541
/gene="cm19"
135. .1541
/gene="cm19"
/codon_start=1
/product="Mltf"
/protein_id="BAA25648.1"
/db_xref="PID:d1026579"
/db_xref="PID:g3061280"
/db_xref="GI:3061280"
/translation="MTRRILRLQRLMEQHQEQRRRQOQKQQAQCFMQRVVSQRP
AINSVPASLPATQVMEVLRKVTLENTFKTHIQQAQKQKQKQIHLITLANKHANO
ALSLPCNPGQDHYMPGTSAPNSPMALTLNSKEGFKFEQSSECEPAL
NTHSRACMQMDVIDDIISLESYNEIIGLMDPALQMANLTPVSGNLIDLGNQSM
PPPLNTNSCPALPNIKRELFESEARALAKERQKDNHLIERRFPNIRIKEL
GLIPKENDMDRNRKGTILKASVDYIRKIQREQRRELNRQKLEHNRHLILAI
QLEMQARAGLSVPSPTGICSPMVRVRIKQEVYLDNCMDLMPHRTDLSCTTIDL
TQDTITSDNNGNTEPTGTYSVPARKSKLELDLMDTLSPGVTDPLLSVSPGAS
KTSRSRVSMEDTHAC"

polyA_signal
polyA_site
2720. .2726
/note="17 A nucleotides"

BASE COUNT 849 a 616 c 606 g 792 t

ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 2863;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2636 TTCGAGAA 2644
|||||
OY 25 ttctgagaa 33

RESULT 160
LOCUS DMEYEL5 2873 bp DNA 22-NOV-1988
DEFINITION Drosophila melanogaster yellow gene 5'-region (yellow locus).
ACCESSION X06481
NID 98834
VERSION X06481.1 GI:8834
KEYWORDS enhancer; yellow gene.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
Drosophilidae; Drosophila.
1 (bases 1 to 2873)
Geyer, P.K. and Corces, V.G.
Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the yellow locus in Drosophila melanogaster
Genes Dev. 1 (9), 996-1004 (1987)
88112841
author numbering refers to yellow transcription start site at pos. +1.

FEATURES
source
location/Qualifiers
1. .2873
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/map_type="2 allele"
1. .2873
/note="yellow"
/gene="y"
/allele=""
/db_xref="FlyBase:FBgn0004034"
1. .1006
/gene="y"
/note="tissue-specific enhancer"
/db_xref="FlyBase:FBgn0004034"
1007. .2176
/gene="y"
/note="tissue-specific enhancer"
/db_xref="FlyBase:FBgn0004034"
2873
/gene="y"
/note="transcription initiation site"
/db_xref="FlyBase:FBgn0004034"

misc_feature
misc_feature
misc_feature

BASE COUNT 910 a 528 c 563 g 872 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 2873;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2352 TTCGAGAA 2360
|||||
Cp 33 ttctcagaa 25

RESULT 161
LOCUS A44216 2899 bp DNA 07-MAR-1997
DEFINITION Sequence 2 from Patent WO9509920.
ACCESSION A44216
NID 92299075
VERSION A44216.1 GI:2299075
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2899)
AUTHORS De, H.P.
TITLE VIRUS RESISTANT OR TOLERANT CELLS
JOURNAL Patent: WO 9509920-A 2 13-APR-1995;
SANDOZ LTD (CH)
Other publication AU 7854394 950501.
COMMENT Location/Qualifiers
1. .2899
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 941 a 533 c 639 g 786 t

ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 2899;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1411 TTCGAGAA 1419
|||||
OY 25 ttctcagaa 33

RESULT 162
ID E11268 standard; DNA; UNC; 2927 bp.
AC E11268;
SV E11268.1
NI d1109605
DT 08-OCT-1997 (Rel. 52, Created)
DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

DE Nucleotide sequence of pTRF.
 KM JP 1996103278-A/3.
 OS unidentified.
 OC unclassified.
 RN [1]
 RA 1-2927
 RA Nakamura A., Tanaka T., Matsuo T., Tanase S., Funatsu M., Eto A.;
 RT "PRODUCTION OF ACTIVE HUMAN ALT";
 RL Patent number JP 1996103278-A/3, 23-APR-1996.
 RL ORIENTAL YEAST CO LTD.
 CC OS None
 CC OC Artificial sequences.
 CC PN JP 1996103278-A/3
 CC PD 23-APR-1996
 CC PF 07-OCT-1994 JP 1994268119
 CC PI NAKAMURA ATSUGO, TANAKA TOSHIO, MATSUO TAKESHI, TANASE SUMIO,
 CC PI FUNATSU MASAHICO, ETO AKIRA
 CC PC C12N15/09,C07H21/04,C07K14/47,C12N1/21,C12N9/10,G01N33/53,
 CC PC G01N33/573,
 CC PC (C12N1/21,C12R1:19),(C12N9/10,C12R1:19);
 CC CC strandedness: Double;
 CC CC topology: Circular;
 CC CC Key
 CC FH Location/Qualifiers
 CC FT source
 CC FT 1. 2927
 CC FT 267.326
 CC FT /note="LTP promoter derived from E.coli K12"
 CC FT CDS
 CC FT 1127..1987
 CC FT /product="beta-lactamase"
 CC FT Location/Qualifiers
 FH source
 FH 1. 2927
 FT /organism="unidentified"
 FT /db_xref="taxon:32644"
 SQ Sequence 2927 BP; 725 A; 736 C; 718 G; 748 T; 0 other;

Query Match 100.0%; Score 9; DB 11; Length 2927;
 Best Local Similarity 100.0%; Pred. No. 8,17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1411 TTCTCAGAA 1419
 Cp 33 ttctcagaa 25

RESULT 163
 LOCUS LV033888 2939 bp mRNA INV 19-OCT-1995
 DEFINITION Lytechinus variegatus USF mRNA, complete cds.
 ACCESSION U33888
 NID 91022699
 VERSION U33888.1 GI:1022699
 KEYWORDS
 SOURCE green urchin.
 ORGANISM Lytechinus variegatus
 Eukaryota; Eukaryota; Eukaryota; Metazoa; Echinodermata;
 Echinozoa; Echinozoa; Echinozoa; Echinozoa; Echinozoa; Echinozoa;
 Toxopneustida; Lytechinus.
 1 (bases 1 to 2939)
 George,J.M., Seid,C.A., Lee,H. and Tomlinson,C.R.
 Two distinct forms of USF do not appear to play a role in the
 inactivation of the Lps1 genes when the extracellular matrix is
 disrupted in the Lytechinus sea urchin embryo
 Unpublished (1995)
 2 (bases 1 to 2939)
 George,J.M.
 Direct Submission
 Submitted (14-AUG-1995) Jenny M. George, Biology, University of
 Houston, Houston, TX 77204-5513, USA
 Location/Qualifiers
 1. 2939
 /organism="Lytechinus variegatus"
 /note="green urchin"

CDS
 /db_xref="taxon:7654"
 883..1854
 /codon_start=1
 /product="USF"
 /protein_id="AAA79690.1"
 /db_xref="pid:91022700"
 /db_xref="gi:1022700"
 /translation="METLYTSYQYIYIFHTYGTSTYEFEGYSGIPNIDKDKLD
 EVANVHTADDDQVTPDPSGSPADNIOYQFRTDNOSQVTVYVNOVCTESRPA
 VYTFPGQQAQIYVMEGSSNSESPTSRQGETFTTTPASAGGCGGASGGEQ
 PGIAAQPPTGSGGQFYVMSPPQVLAQASQRTIAPRTHQFTKVSATVDERRRAT
 HNEVRRRRDKINMIVLKSRIIPDCNIDHSKQSGGILSKTCDYIQELRQSNTRM
 AESLDTERTLSVIDIMROQLLELGENALLRAQLQAGIDPPTGSSN"

BASE COUNT 909 a 567 c 601 g 846 t 16 others

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 2939;
 Best Local Similarity 100.0%; Pred. No. 8,17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2841 TTCTCAGAA 2849
 Cp 33 ttctcagaa 25

RESULT 164
 LOCUS OC021155 2948 bp mRNA MAN 01-MAY-1995
 DEFINITION Oryctolagus cuniculus sarcolemmal associated protein (SLAP1) mRNA,
 complete cds.
 ACCESSION U21155
 NID 9790239
 VERSION U21155.1 GI:790239
 KEYWORDS
 SOURCE European rabbit.
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Eukaryota; Eukaryota; Metazoa; Chordata;
 Vertebrata; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 1 (bases 1 to 2948)
 Wigle,J.T., Demchishyn,L., Sattar,S., Pratt,M. and Tuana,B.S.
 Molecular Cloning of a Novel Family of Sarcolemmal Associated
 Colled-Coll Proteins: Protein-Protein Interactions at the Cell
 Membrane
 Unpublished
 2 (bases 1 to 2948)
 Wigle,J.T.
 Direct Submission
 Submitted (16-FEB-1995) Jeffrey T. Wigle, University of Ottawa,
 Pharmacology, 451 Smyth Rd, Ottawa, Ontario, Canada, K1H 8M5
 Location/Qualifiers
 1. 2948
 /organism="Oryctolagus cuniculus"
 /db_xref="taxon:9986"
 /clone="SLAP1"
 /sex="male"
 /tissue_type="heart"
 /dev_stage="adult"
 189..1157
 /gene="SLAP1"
 189..1157
 /gene="SLAP1"
 /evidence-experimental
 /product="sarcolemmal associated protein"
 /protein_id="AAA65598.1"
 /db_xref="pid:9790240"
 /db_xref="gi:790240"
 /translation="WDEQDLNESLAKVSLKALLLEERKAYRNOVESSKOIQVLAQ
 LQRLMDLENREKDNEDITSRELSARBEIILLHQAARASERDPTIASIQEEL
 KVRLELRERKRASEYEVSLSQSSFOCCOCEQDOKEATRLQGELEKREVN
 VLETRCHLKKENYLSSELRQREKELNSQKQSELTSDLSIOMTRKELENGWGL
 KQHRDRSDADKILSKRAENAKQVQKREKQVLSLKIKFMTDEKQSIDELK
 QCKDKLKLQEKGNKNNKPPMPPHLLAVYATVLYVPGIARRSP"
 630..716

misc_feature

```

/gene="SLAP1"
/product="leucine zipper-like motif"
misc_feature
777..842
/gene="SLAP1"
/product="leucine zipper-like motif"
1080..1130
/gene="SLAP1"
/function="anchors protein to membrane"
/product="hydrophobic carboxyl tail"
2923..2928
polya_signal
BASE COUNT 965 a 479 c 609 g 895 t
ORIGIN
Query Match 100.0%; Score 9; DB 23; Length 2948;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 755 TTCTCAGAA 763
|||||
Cp 33 ttctcagaa 25

RESULT 165
LOCUS AB004649 2952 bp mRNA VRT 05-FEB-1999
DEFINITION Gallus gallus mRNA for CENP-C, complete cds.
ACCESSION AB004649
VERSION g2749772
KEYWORDS alternative splicing; CENP-C; centromere protein C.
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Vertebrata; Archosauria; Aves;
Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
1 (bases 1 to 2952)
Fukagawa, T.
Direct Submission
Submitted (10-JUN-1997) to the DDBJ/EMBL/GenBank databases. Tatsuo
Fukagawa, University of Oxford, Department of Biochemistry, South
Parks Road, Oxford OX1 3OU, UK (E-mail: tfukagaw@bioch.ox.ac.uk,
Tel:44-1865-275225, Fax:44-1865-275259)
2 (sites)
Fukagawa, T. and Brown, W.R.
Efficient conditional mutation of the vertebrate CENP-C gene
Hum. Mol. Genet. 6 (13), 2301-2308 (1997)
98030544
FEATURES
Source
1..2952
Location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="CENP-C_5_3"
<1..>2952
/note="alternative splicing: see also AB004650"
31..2649
/standard_name="centromere protein C"
/codon_start=1
/product="CENP-C"
/protein_id="BAA24109.1"
/db_xref="PID:d1025016"
/db_xref="PID:g2749773"
/db_xref="GI:2749773"
CDS
MRNA
CDS

```

```

polya_signal
BASE COUNT 976 a 588 c 683 g 705 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 2952;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1842 TTCTCAGAA 1850
|||||
Cy 25 ttctcagaa 33

RESULT 166
LOCUS MTDNAFLPF 2971 bp DNA BCT 07-MAY-1996
DEFINITION M.thermoautotrophicum flpF gene.
ACCESSION X87969
VERSION X87969.1 GI:871455
KEYWORDS alpha subunit; flpF gene; formate dehydrogenase.
SOURCE Methanobacterium thermoautotrophicum
ORGANISM Methanobacterium thermoautotrophicum
Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
Methanobacterium.
1 (bases 1 to 2971)
Hochheimer, A., Schmitt, R.A., Thauer, R.K. and Hedderich, R.
The tungsten formylmethanofuran dehydrogenase from Methanobacterium
thermoautotrophicum contains sequence motifs characteristic for
enzymes containing molybdopterin dinucleotide
Eur. J. Biochem. 234 (3), 910-920 (1995)
96163477
REFERENCE
JOURNAL
MEDLINE
2 (bases 1 to 2971)
Hochheimer, A.
Direct Submission
Submitted (15-JUN-1995) A. Hochheimer, Max-Planck-Inst. fuer
Terrestrische, Mikrobiologie, Karl-von-Frisch-Str., 35043 Marburg,
FRG
FEATURES
Source
1..2971
Location/Qualifiers
/organism="Methanobacterium thermoautotrophicum"
/strain="Marburg/DSM 2133"
/db_xref="taxon:2166"
/clone_11b="EMBL4"
134..2797
/gene="flpF"
/gene="flpF"
134..2797
/gene="flpF"
/transl_start=1
/codon_start=1
/product="putative alpha subunit of formate
dehydrogenase"
/protein_id="CAA61208.1"
/db_xref="PID:g871456"
/db_xref="GI:871456"
/db_xref="SPTREMBL:O50743"
/transl_except="MKGYKFRIDGRGEVEAARGMTVLEALANGIYIPNLCPREGIEP
FGGCRCLIVENNGRGLVYACETPAEDSGEYSEISSEIRNRRTLSLIIDHSRDCA
CPASGCRDDELISYLVNSDGLDELRRLPELISGIDVESNPFRLNHDKICIGICRY
CRGAGAAYFAIRGHDTRIATMDRIDLSDSCVSCGEYCACPVGAHLRTERPSTE
VRVYDPEFEVEMDEAI SVYAERLSYTEGEFAVAASACTEENYVLOKFRVAVNG
SGNIDHCAFLCHAPSILGLMSLSGAMTNSISELAAGC ILAVGNTPETHPTSVR
VITALRBRATLVVDPKTRPLSLADIHLONRPGSDIPLMAMCRFIEGLDSEI
DSKTRKEDRDVMAALDDEVERITGVYKQDVRRAIMYASNSPASIYSMTITAV
NGGVNVALSNLVLITGINSIKSAGINPLRGQNVGACDMGLPDLIPYQIGENA
GKSEKSGVSIIPAGLTLPEMPAARGKTRCYINGENPLLSEPIERTREALBIE
FLVYDCLLETELADVYLPAASFKEKDDTFNTRRVLQKRALDAPDADLPMQI
ISMVAGMGECHDEYASRIPEIRIELVSYGISHERLKSGLQMPCTSEBDATG
YLSEEPPTGASFLIPYQGRFVPEEYPLVLTGRNLVYHTSMTRVRELSF
SDHEELMANPADASSMGIRGDIYEVTSENGSLRVAGVYDDEVMEGVYRTTFIAPAD

```

BASE COUNT 748 a 699 c 844 g 680 t
 ORIGIN ANVLGDRDPFGMPGLKFTPVRCYQVLR"

Query Match 100.0%; Score 9; DB 17; Length 2971;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2127 TTCTGAGAA 2135
 |||||
 25 ttctgagaa 33

RESULT 167
 LOCUS 150971 2989 bp DNA PAT 12-AUG-1997
 DEFINITION Sequence 1 from patent US 5643776.
 ACCESSION 150971
 NID 92472674
 VERSION 150971.1 GI:2472674
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2989)
 AUTHORS Hammock,B.D., Hanzlik,T.N., Harshman,L.G., Bonning,B.C. and Ward,V.K.
 TITLE Insect diagnostic and control compositions
 JOURNAL Patent: US 5643776-A 1 01-JUL-1997;
 FEATURES
 source Location/Qualifiers
 1..2989
 /organism="unknown"

BASE COUNT 902 a 625 c 636 g 826 t

Query Match 100.0%; Score 9; DB 25; Length 2989;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1477 TTCTGAGAA 1485
 |||||
 25 ttctgagaa 33

RESULT 168
 LOCUS CCHSP70PR 3005 bp DNA INV 17-MAR-1997
 DEFINITION C.capitata gene encoding for HSP70 protein.
 ACCESSION Y08955
 NID 91894775
 VERSION Y08955.1 GI:1894775
 KEYWORDS HSP70 protein.
 SOURCE Mediterranean fruit fly.
 ORGANISM Ceratitis capitata
 Eulayota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Diptera; Brachycera; Muscomorpha; Tephritidae;
 Tephritidae; Ceratitis.
 1 (bases 1 to 3005)
 Mintzas,A.C.
 Mintzas,A.C.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1996) A.C. Mintzas, University of Patras,
 Department of Biology, Patras, 26500, GREECE
 REFERENCE 2 (bases 1 to 3005)
 Mintzas,A.C.
 JOURNAL Unpublished
 FEATURES
 source Location/Qualifiers
 1..3005
 /organism="Ceratitis capitata"
 /db_xref="taxon:7213"
 448..455
 TATA_signal
 679..2595
 /codon_start=1
 /product="HSP70 protein"
 /protein_id="CAA70153.1"
 /db_xref="PID:e303885"

/db_xref="PID:q1894776"
 /db_xref="GI:1894776"
 /db_xref="SWISS-Prot:P91902"
 /translation="MAIGIDLTTTSCFEVFGHKEVETIANDOGNTPTSYAFIDS
 ERLGDAKNQVANNPNTVFDARLIGRYIDPKIMEDVKNHPEKVSOGKPKISV
 EYKERNQFAPEEISSMVLTKMETAVIIIGTVTDVAVIVPVPFNSOQAKRDAR
 IAGLVAVRIINEPFAALAYGLDKENGNVLIIFDGGTFDVSILITIDGSLFEVR
 ATAGDTLHGDFDNRVLSHAEEKRKKKDLRENPALRLRRTAAERAKRTISSYE
 ATIRIDALPGIDLYTKVSRARPEELCADLPROLPEVEKALANDAKDKNOIHYVYV
 GGSTRIRKVRRLLOSFFCGKSLNINSPDRAVANGAQAIALISGDKSTETIOVLVD
 VAPLSLGIETAGVMAKIERNCRIPEKQOTSTYSIDNPGVNOIYVEERWTKLN
 NRLETGFDLSGIPAPRGVPOIEVTFVDANGNNLNYSKEMSSGNANITIKDKGKL
 SOAEIDRVNENAGRYAEDEDERORNKIARNNLESYLAVQAQMTTLVDKLSERKSEY
 TKACDDITIKMLDNRILADKEEYEDKMTLRLKLCPIWTKLHSGAGAGASCQQAAG
 FNGHTEPTVEVD"

BASE COUNT 961 a 615 c 701 g 728 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 3005;
 Best Local Similarity 100.0%; Pred. No. 8.17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 TTCTGAGAA 107
 |||||
 25 ttctgagaa 33

RESULT 169
 LOCUS OCT21156 3020 bp mRNA MAM 01-MAY-1995
 DEFINITION Oryctolagus cuniculus sarcolemmal associated protein-2 mRNA,
 complete cds.
 ACCESSION U21156
 NID 9790235
 VERSION U21156.1 GI:790235
 KEYWORDS
 SOURCE European rabbit.
 ORGANISM Oryctolagus cuniculus
 Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 1 (bases 1 to 3020)
 Wigle,J.T., Demchishyn,L., Sattar,S., Pratt,M. and Thana,B.S.
 TITLE Molecular cloning of a novel family of sarcolemmal associated
 1 (bases 1 to 3020)
 colled-coll proteins: protein-protein interactions at the cell
 membrane
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3020)
 AUTHORS Wigle,J.T.
 TITLE Direct Submission
 JOURNAL Submitted (15-FEB-1995) Jeffrey T. Wigle, Pharmacology, University
 of Ottawa, 451 Smyth Road, Ottawa, Ontario K1H 8M5, Canada
 FEATURES
 source Location/Qualifiers
 1..3020
 /organism="Oryctolagus cuniculus"
 /db_xref="taxon:9986"
 /clone="SLAP2"
 /sex="male"
 /tissue_type="heart"
 /dev_stage="adult"
 21..1229
 /codon_start=1
 /evidence="experimental"
 /product="sarcolemmal associated protein-2"
 /protein_id="AA65596.1"
 /db_xref="PID:9790236"
 /db_xref="GI:790236"
 /translation="MEKEDELQAKIFALADNDFTNERITLALOEKILVEGHITKVE
 ETKLAKENQARAKESDSDTLSPSEKSSDDTDDAODDEDLWESLAKVSLKALLEE
 ERKAYRNQVESSKOIQLAQQLRLMDIENLEEDKNEITSTRDELLARPEIILL
 HOAEKASERDQDIASLOELKVRALREMRKAASEYKEVTSLOSGSFOLCOQCE
 DQKEEATRIQGELEKRLKEMNVLETGCHSLKKKENYLLSSELOROEKLNSOKSLE
 LTPDSLITQTRKLEENOMSKLNQHRDSDADLKILSKAKENAKQVQKREKTOYVL
 SELKLFEMTEQEKQSTIDELKQCKDKNLQEGNNKPKMPMPMLALAVATATL
 VYPLARASP"

```

misc_feature      702..788
                  /note="encodes leucine zipper-like motif"
misc_feature      849..914
                  /note="encodes leucine zipper-like motif"
misc_feature      1152..1202
                  /note="encodes a hydrophobic carboxyl tail"
                  /function="anchors protein to membrane"
POLYA.signal      2995..3000
BASE COUNT      1006 a 491 c 623 g 900 t
ORIGIN
Query Match      100.0%; Score 9; DB 23; Length 3020;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      827 TTCTCAGAA 835
        |||||||||
        33 ttctcagaa 25

RESULT 170
LOCUS      PRGCCDB      3024 bp      DNA      BCT      07-MAY-1996
DEFINITION Plasmid KIL18 (from E.coli) cytotoxic protein (ccdb) gene, complete
ACCESSION 127083
NID        9548209
VERSION    127083.1 GI:548209
KEYWORDS   ccdb gene; cytotoxic protein.
SOURCE     plasmid KIL18 DNA.
ORGANISM   plasmid KIL18
REFERENCE   1 (bases 1 to 3024)
AUTHORS    Bernard, P., Gabant, P., Bahassi, E.M. and Couturier, M.
TITLE      Positive-selection vectors using the F plasmid ccdb killer gene
JOURNAL     Gene 148 (1), 71-74 (1994)
MEDLINE    95011662
FEATURES
  source
    1..3024
    /organism="Plasmid KIL18"
    /plasmid="Plasmid KIL18"
    /specific_host="K-12"
    /db_xref="taxon:37150"
  gene
    1..381
    /gene="ccdb"
    /gene="ccdb"
    /codon_start=1
    /transl_table=11
    /product="cytotoxic protein"
    /protein_id="AA98268.1"
    /db_xref="PID:9548210"
    /db_xref="GI:548210"
    /translation="MTMTSSVSPGDPLESTCRHSLSLQFVYTYKRSRYRLFDV
VOSDIIDTGRMVIVPLASRLISDKVSRRLYVYVIGDSRMTTDMASVYSVIG
EEVADISHRENDKNAINLMFNGT"
BASE COUNT      743 a 759 c 770 g 752 t
ORIGIN
Query Match      100.0%; Score 9; DB 17; Length 3024;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1292 TTCTCAGAA 1300
        |||||||||
        33 ttctcagaa 25

RESULT 171
LOCUS      AF090111      3030 bp      mRNA      VRT      10-JAN-1999
DEFINITION Xenopus laevis phospholipase C-gamma-1a mRNA, partial cds.
ACCESSION  AF090111
NID        54138836
VERSION    AF090111.1 GI:4138836

```

```

KEYWORDS
SOURCE      African clawed frog.
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia;
            Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus.
REFERENCE   1 (bases 1 to 3030)
AUTHORS     Kung, H.-F., Kim, D., Huang, Y.K. and Lin, M.C.
TITLE       Molecular cloning of two different forms of Xenopus phospholipase
            C-gamma-1
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 3030)
AUTHORS     Kung, H.-F., Kim, D., Huang, Y.K. and Lin, M.C.
TITLE       Direct Submission
JOURNAL     Submitted (02-SEP-1998) Laboratory of Biochemical Physiology,
            Division of Basic Science, National Cancer Institute-Frederick
            Cancer Research and Development Center, Frederick, MD 21702-1201,
            USA
FEATURES
  source
    1..3030
    /organism="Xenopus laevis"
    /db_xref="taxon:8355"
    <1..>3030
    /codon_start=1
    /product="phospholipase C-gamma-1a"
    /protein_id="AA03594.1"
    /db_xref="PID:9413837"
    /db_xref="GI:4138837"
    /translation="QRSREINVAQFAMLYRLSLMDAOKTIPITETEGEGERTKSL
EDPQKFLDYOKDMWATDLHWYQSEMFHLEHPELEIEPEFTLEEVTFLEKRENTY
WDLERDYCPQNMNDPLSHYWSLSHNTYLGDDSSSESLAARCLRMGCRCEILD
CWDGPDGMPAITHGHTLTTKIFSEVSTIKENAVTSEYVYILSIEDHCSIAQRNA
ADHFKRVFQDMLTLPVDVNAADGLPSFQMLKRLIKRLKLAEGSTIEIFSSVYS
NDISINSINGILYLODPINHEMYPHFVLTLSKTIYSELTGNANDDEEOKASNS
SELHSAEKWFHKLACGRDGRHIAERLLTDCIETGADGSEFLVASESTFVADYTLSE
WNGVQCHRIHSROGASPKFEFLDNDVLFSLVALLHYOQMPLRNCEFMRLTEPV
POTNAHSEKWHASLITRGOAHMTMRVPRGAFIVRRKRSQNSVAISFRAGRTKHC
RYIQGQSVYLGSSSEFDSLVDLISYERHPIYKRLKLYPINEETLEKIGTPDDYGA
LYEGRNPGEYVIANPMTFKCSVRALFYKAREDELFTKNTIIONVEKOBGMWRG
DCGGRKQMFEPANMYMEIIFSPPEPEPQHLDENSLDGLDGLGVADVPSCHIVPQDV
ENGRPEVFTITGPQNLRYPLDYAADTLTDMQDWIRKIEAQTADARLTBEKIMERRK
KIALSELVYICRVPEDEKIEKACQYRMSFPEKAKYVKKKGGKFLQYNR
ROLSRIYPKGORLDSNSYDPLTMWICGSLVALNFOPTDKMONMNALEFQSGRGYV
FQPNCRDEMDPDPKLTLLTETITISIEIGARLPRKIGIGVYCPFEVYVCGTEYD
NAKQTEFVNDGLNVPVPOKTFSSVYVNAPEAFIRFYVYEDNFSQDNFLAQSFFV
RGLKTYRAIPLKNNTYEDLEALLILIDIKTENGEL"
BASE COUNT      877 a 686 c 728 g 739 t
ORIGIN
Query Match      100.0%; Score 9; DB 24; Length 3030;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      547 TTCTCAGAA 555
        |||||||||
        33 ttctcagaa 25

RESULT 172
LOCUS      A13388      3104 bp      DNA      PAT      03-JAN-1994
DEFINITION PSP18 DNA sequence.
ACCESSION  A13388
NID        9491651
VERSION    A13388.1 GI:491651
KEYWORDS   .
SOURCE     synthetic construct.
ORGANISM   artificial sequence.
REFERENCE   1 (bases 1 to 3104)
AUTHORS     Hoeltke, H.J., Seibl, R., Schmitz, G., Schoeler, H.R., Kessler, C. and
            Mattes, R.
TITLE       Method for the detection of nucleic acids
JOURNAL     Patent: EP 0324474-A 4 19-JUL-1989;
            BOEHRINGER MANNHEIM GMBH

```


FEATURES
source
Location/Qualifiers
1..3104
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 768 a 800 c 802 g 734 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 3104;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1903 TTCTGAGAA 1911
|||||
25 ttctgagaa 33

RESULT 173
LOCUS BTBVPAP 3106 bp mRNA MAM 18-DEC-1991
DEFINITION B. taurus mRNA for bovine vacuolar ATPase subunit A.
ACCESSION X58986
NID 9162
VERSION X58986.1 GI:162
KEYWORDS dissociable hydrophilic catalytic complex (V1); H⁺-ATPase A
subunit; vacuolar H⁺-ATPase.
SOURCE Bos taurus.
ORGANISM Bos taurus.
Eukaryota; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 3106)

REFERENCE
AUTHORS Dean, G.E.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1991) G.E. Dean, Univ. of Cincinnati College of
Medicine, Dept. of Mol. Genetics Biochem and Micro, 231 Bethesda
Avenue, Cincinnati Ohio 45267-0524, USA
2 (bases 1 to 3106)
Pan, Y.X., Xu, J., Strasser, J.E., Howell, M., and Dean, G.E.
TITLE Structure and expression of subunit A from the bovine chromaffin
cell vacuolar ATPase
JOURNAL FEBS Lett. 293 (1-2), 89-92 (1991)
MEDLINE 92070363
FEATURES
source
Location/Qualifiers
1..3106
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="adrenal medulla"
/cell_type="chromaffin cell"
/clone_1lb="Lambda ZAP chromaffin cell cDNA library"
/clone="Bluescript SK(-) plasmid"
1..48
1..3106
49..1905
/EC_number="3.6.1.35"
/note="H⁺-ATPase subunit A"
/product="H⁺-transporting ATPase"
/protein_id="CA41276.1"
/db_xref="pid:9163"
/db_xref="gi:163"
/db_xref="swiss-prot:P31404"
/translation="MMDSKLPKRIEDKESSTFGYHGVSGPVTACDAGAAATLVE
RVGSELVGEIIRLEGDMAITQVEETSGVSDVLRKGLSVELPGIAGALFEDG
IORPLDISSTOSISYIPRGVNVASLSRVKMDPFCRLKRGSHITGADIGIYEN
SLIKKIMLPNNRGTVTIAPRGVNTSDVLTLEFGIKERFSMVQVMPVROYRPV
TEKLPANHPILTGQVLDALFPCVGGTTATGACGCKTYSQSLSKSNEDYITIV
GCGERNGESELVRDPELTMDGKVESIMKRLVANTSMNPVAAEDASTYITITL
SEYFRDMGYHVSMAADSTSRMAELREISGLAEMPADSGYPAYIGALASTYERAG
VKCLGNPERGVSIVGAVSPGSDPVSATIGIVQEWGLDRLKLAQRKHPSTV
WLISYKTRALDEYDKHETFEVPLRTAKELIQEEDLAEIVOLVGKASIAETDKI
TLEVAKLKDKDLONGYTPYRPFCTKVTGMLSMETAFYDMARAVETTRQSONKI
TWSITREHMGELTKLSMKFKDPVKDEAKIKADYAOULEDMQNAFSLSD"

3' UTR
BASE COUNT 860 a 625 c 708 g 913 t
ORIGIN

Query Match 100.0%; Score 9; DB 23; Length 3106;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2732 TTCTGAGAA 2740
|||||
25 ttctgagaa 33

RESULT 174
LOCUS ECAJ1614 3128 bp DNA circular BCT 16-DEC-1998
DEFINITION Escherichia coli plasmid pcaps.
ACCESSION AJ001614
NID 92769261
VERSION AJ001614.1 GI:2769261
KEYWORDS ampicillin resistance; beta-lactamase; bla gene; catabolite gene
activator protein; crp gene; cyclic amp receptor protein.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
1 (bases 1 to 3128)

REFERENCE
AUTHORS Schleper, D.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1997) Schleper D., Institute of Genetics,
University of Cologne, Weyertal 121, Cologne, 50931, GERMANY
2 (bases 1 to 3128)
Schleper, D., von Wilcken-Bergmann, B., Schmidt, M., Sobek, H. and
Muller-Hill, B.
TITLE A positive selection vector for cloning of long polymerase chain
reaction fragments based on a lethal mutant of the crp gene of
Escherichia coli
JOURNAL Anal. Biochem. 257 (2), 203-209 (1998)
MEDLINE 98189310
FEATURES
source
Location/Qualifiers
1..3128
/organism="Escherichia coli"
/plasmid="pcaps"
/db_xref="taxon:562"
complement(315..947)
/gene="crp"
complement(315..947)
/note="artificially mutated crp; catabolite gene activator
protein"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="cyclic AMP receptor protein"
/protein_id="CA04867.1"
/db_xref="pid:e1231197"
/db_xref="pid:92769262"
/db_xref="gi:2769262"
/db_xref="sptrembl:O52983"
/translation="MVLGKPTQDPTLEWFLSHCHIKPKSKSTLIHQEKAETIYIV
KGSVAVLIKDEGKEMITSYLNOGPFIEGLPFGGOSAMVRAKTACVAAEISYRK
PROLIVNPDIIMRLSOMARRIOVTSKVNLAFLDVGRIATOTILNLAQOPAMTH
PDGMQIKITRQIEIGIYCSKOTYGRILKMLNDONLISAHKTIYVIGTR"

rep_origin
gene
CDS
1313
/direction="BOTH"
complement(2072..2932)
/gene="bla"
complement(2072..2932)
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="CA04868.1"
/db_xref="pid:e1231198"
/db_xref="pid:92769263"
/db_xref="gi:2769263"
/db_xref="sptrembl:Q00626"

```

BASE COUNT      799 a      773 c      763 g      793 t
ORIGIN

```

Query Match	100.0%	Score 9	DB 17	Length 3128
Best Local Similarity	100.0%	Pred. No. 8.17e+03		
Matches	9	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0

Db	2640	TTCTGAGAA	2648
QY	25	ttctgagaa	33

RESULT	175	VRT	28-FEB-1997
LOCUS	GCCHIRA	3144 bp	mRNA
DEFINITION	G.gallus mRNA for Chlra protein.		
ACCESSION	X93375		
NID	01890568		
VERSION	X93375.1	GI:1890568	
KEYWORDS	Chlra protein; DGeorge syndrome.		
SOURCE	Chicken.		
ORGANISM	Gallus gallus		

REFERENCE	1 (bases 1 to 3144)
AUTHORS	Roberts,C., Daw,S., Halford,S., Whiting,J.A. and Scambler,P.J
TITLE	Cloning and developmental expression analysis of chick hira (Chir), a candidate gene for Digeorge Syndrome

JOURNAL	Mamm. Genome in press
REFERENCE	2 (bases 1 to 3144)
AUTHORS	Scambler, P.J.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-1996) P. J. Scambler, Institute of Child Health, 30 Guilford St, Room 214, London, WC1N 1EH, UK
FEATURES	Location/Qualifiers
SOURCE	1..3144

```

/organism="Gallus gallus"
/db_xref="taxon:9031"
/dev_stage="12"
61. .3117
CDS

```

BASE COUNT	900 a	711 c	767 g	764 t	2 others
ORIGIN					

Query Match	100.0%	Score 9;	DB 24;	Length 3144;
Best Local Similarity	100.0%	Pred. No. 8.17e+03;		
Matches	9;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Db	434	TTTCGAGAA	442	
Cp	33	tttcgagaa	25	

RESULT	176
LOCUS	OVGSR2G 3150 bp DNA INV
DEFINITION	X. volutus GSR gene for glutathione transferase.
ACCESSION	X83915
NID	G975210
VERSION	X83915.1 GI:975210
KEYWORDS	glutathione transferase.
SOURCE	Onchocerca volutus.
ORGANISM	Onchocerca volutus

REFERENCE AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 3150) Liebau, E., Wlodenburg, G., Brophy, P. M., Walter, R. D. and Henkle-Duhren, K.	Biochemical analysis, gene structure and localization of the 24 kDa glutathione S-transferase from <i>Onchocerca volvulus</i>	Mol. Biochem. Parasitol.	81, 270-270 (1996)
2 (bases 1 to 3150)			

TITLE	Direct Submission
JOURNAL	Submitted (13-VAN-1995) E. Liebau, Bernhard-Nocht-Inst. for Tropical Med., Dept of Biochemistry, Bernhard-Nocht-Strasse 74, 20359 Hamburg, FRG
REFERENCE	3 (bases 1 to 3150)
AUTHORS	Liebau, E., Widenburg, G., Brophy, P. M., Walter, R. D. and Henkle-Dührsen, K.
TITLE	Biochemical analysis, gene structure and localization of the 24 kDa glutathione S-transferase from <i>Onchocerca volvulus</i>
JOURNAL	Mol. Biochem. Parasitol. 80 (1), 27-39 (1996)
MEDLINE	97039668

```

source
1..3150
/organism="Onchocerca volvulus"
/db_xref="taxon:6282"
/dev_stage="adult"
/clone_lib="lambda-flx"
724..757
exon

```

intron
exon
intron
exon
intron
exon

intron /number=4 1671..2008
exon /number=5 2009..2119
intron /number=5 2120..2265
exon /number=5 2266..2394
intron /number=6 2395..2534
exon /number=6 2535..2594
BASE COUNT 998 a 451 c 488 g 1213 t
ORIGIN
Query Match 100.0% Score 9: DB 21: Length 3150:
Best Local Similarity 100.0% Pred. No. 8.17e+03:
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 843 TTCTCAGAA 851
CP 33 ttctcagaa 25
RESULT 177
LOCUS SP076218 3170 bp DNA BCT 12-MAR-1997
DEFINITION Streptococcus pneumoniae competence stimulating peptide precursor
regulator homolog (comC) genes, complete cds.
ACCESSION U76218
NID 91666904
VERSION U76218.1 GI:1666904
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 3170)
AUTHORS Cheng O., Campbell, E.A., Naughton, A.M., Johnson, S. and Measure, H.R.
TITLE The com locus controls genetic transformation in Streptococcus pneumoniae
JOURNAL Mol. Microbiol. 23 (4), 683-692 (1997)
MEDLINE 97206147
REFERENCE 2 (bases 1 to 3170)
AUTHORS Cheng O., Campbell, E.A., Naughton, A.M., Johnson, S. and Measure, H.R.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1996) Laboratory of Molecular Infections
diseases, The Rockefeller University, 1275 York Ave., New York, NY
10021, USA
FEATURES
source Location/Qualifiers
1. 3170
/organism="Streptococcus pneumoniae"
/strain="R6"
/note="uncapsulated"
/db_xref="taxon:1313"
50..529
/note="ORF1"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC44894.1"
/db_xref="PID:91666905"
/db_xref="GI:1666905"
/translation="MKIKVTVYKGLKEKYLKDGIAEYGRISRFKPFEMIELSDEKTP
DKASEENQKILEITGQRLSKLADRDVYIVATIGKTFPSEFSKQLEBSIKGFST
LFTIGSLGLSSSVKNRANLSVSGRLTLPQMLRLVLEQIYNAFTIQGFPYHK"
811..882
/gene="comC"
811..936
/gene="comC"
811..936
/gene="comC"
/function="competence stimulating peptide precursor"

mat_peptide
/codon_start=1
/transl_table=11
/product="comC"
/protein_id="AAC44895.1"
/db_xref="PID:91666906"
/db_xref="GI:1666906"
/translation="MKNTVKLEQFVALKEDLQKIGGEMRLSKFRFDILRRK"
883..933
/gene="comC"
/function="competence stimulating peptide"
/product="comC"
957..2282
/gene="comC"
957..2282
/gene="comC"
/note="part of the two-component regulatory system;
histidine kinase homolog"
/codon_start=1
/transl_table=11
/product="comD"
/protein_id="AAC44896.1"
/db_xref="PID:91666907"
/db_xref="GI:1666907"
/translation="MDLEFGFVIVHFLIISHSYHICKQINKRELFGVAYTLLE
IVFDPFLIYLDIGIERFPLGLYSYFPMKQYERDGLFLSLSLYESTHNF
LSYTFSSITGDNFVLYQYHFPFVTVTLKIIYFHLDELAFEDLYPLK
VEPALLIHYSFVSDWSTIKHNSGSLISYFSLTFPAMNSHKYOMEREA
LKQKRPDRNIRITQNEIYVGIKNEIRGRHRYAGMYSKMAIDSGNQLDRIYNEY
LVKANKHLSKRYTFDNLNIEDSLNVAQSVIYARNKQVETLEVKDITLPLIE
LIDLVRINSVLNNNAVEGSADSYKRMVAVIKMETVIVIONSCKMTWPSGDLFA
LGFSYKGRNRCVGVNNVKELDKYNNIILETEMGSGTFROIIRKREFE"
2279..3031
/gene="comE"
2279..3031
/gene="comE"
/note="part of two-component regulatory system; response
regulator homolog"
/codon_start=1
/transl_table=11
/product="comE"
/protein_id="AAC44897.1"
/db_xref="PID:91666908"
/db_xref="GI:1666908"
/translation="MKVILIEDVIEHQVRLRIIDLSKESNIPISYKTKGVREFEE
YIENDEVQYLFIDIDIGIEKGEFVQLIRHNPYAIIVITSRSFALTYKQY
SALDPVDKDIDEMEKRIEONIFYTSKMLENDVDVYDPYNTKGNDKIPIYDIY
IETGYSKLEIKIRKPAKRYFGYGMTDIOEDKHQRYSPHSKSPLVNIGNRIEDRK
NLEIVFYEDHRCPISRKLRKLDILEKKSOK"
BASE COUNT 1093 a 425 c 603 g 1049 t
ORIGIN
Query Match 100.0% Score 9: DB 18: Length 3170:
Best Local Similarity 100.0% Pred. No. 8.17e+03:
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 299 TTCTCAGAA 307
CP 33 ttctcagaa 25
RESULT 178
LOCUS I82816 3225 bp DNA PAT 20-MAR-1998
DEFINITION Sequence 13 from patent US 5712381.
ACCESSION I82816
NID 93211113
VERSION I82816.1 GI:3211113
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3225)
AUTHORS Lin, L., Chen, J., Schlievelia, A.R. and Graham, J.
TITLE MADD, a TNF receptor death domain ligand protein

JOURNAL Patent: US 5712381-A 13 27-JAN-1998;
FEATURES
Source 1..3225
Location/Qualifiers

BASE COUNT 703 a 874 c 1030 g 618 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 3225;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 536 TTCGAGAA 544
|||||||
Cp 33 ttctcagaa 25

RESULT 179
LOCUS XIGATA6 3299 bp mRNA VRT 07-APR-1997
DEFINITION X.laevs mRNA for GATA-6 protein.
ACCESSION Y08865
NID Y08865.1 GI:1621280
VERSION GATA-6 gene; GATA-6 protein.
KEYWORDS African clawed frog.
SOURCE Xenopus laevis
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia;
Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 3299)
AUTHORS Gove,C., Wainmley,M., Nijjar,S., Bertwistle,D., Gulle,M.,
Partington,G., Bomford,A. and Patient,R.
Over-expression of GATA-6 in Xenopus embryos blocks differentiation
of heart precursors
EMBO J. 16 (2), 355-368 (1997)
JOURNAL 97180937
MEDLINE Erratum: [[published erratum appears in EMBO J 1997 Apr
1;16(7):1806-7]]
REMARK 2 (bases 1 to 3299)
REFERENCE 2 (bases 1 to 3299)
AUTHORS Gove,C.D.
TITLE Direct Submision
JOURNAL Submitted (17-OCT-1996) C.D. Gove, Kings College, University of
London, The Randall Institute, 26-29 Drury Lane, London, WC2B 5RL,
UK

FEATURES
Source Location/Qualifiers

1..3299
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/dev_stage="adult"
/tissue_type="liver"
220..1395
/gene="gata-6"
220..1395
/gene="gata-6"
/codon_start=1
/product="GATA-6 protein"
/protein_id="CAI70088.1"
/db_xref="PID:g1621281"
/db_xref="GI:1621281"
/db_xref="SWISS-PROT:P70005"

BASE COUNT 942 a 813 c 661 g 883 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 3299;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2308 TTCGAGAA 2316
|||||||
Cy 25 ttctcagaa 33

RESULT 180
LOCUS BMOCHOC1 3317 bp DNA INV 05-MAY-1994
DEFINITION silkmooth (B.mor1) chorion CB early protein (5H4) gene, complete
cds, clones E1 and E2.
ACCESSION M19075
NID M19075.1 GI:155973
VERSION M19075.1
KEYWORDS chorion; eggshell protein.
SEGMENT 1 of 3
SOURCE silkmooth (B.mor1) DNA, clones E1 and E2.
ORGANISM Bombyx mori
Eukaryota; Eukaryota; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Lepidoptera; Bombycoidea;
Bombycidae; Bombyx.
1 (bases 1 to 3317)
AUTHORS Hübner,B.L., Burke,W.D., Lecanidou,R., Rodakis,G.C. and
Eickbush,T.H.
Organization and expression of three genes from the silkmooth early
chorion locus
Dev. Biol. 125, 423-431 (1988)
JOURNAL 88112521
MEDLINE

FEATURES
Source Location/Qualifiers

1..3317
/organism="Bombyx mori"
/db_xref="taxon:7091"
719..804
/gene="5H4"
/number=1
join(757..804,2208..2684)
/gene="5H4"
/codon_start=1
/product="chorion protein"
/protein_id="AAA27830.1"
/db_xref="PID:g155977"
/db_xref="GI:155977"
/translation="MTIVLVLCASALFVLAFCGLGRDPYIFGAGSGMGYDA
ISPYDGLGYGPYSAAGFISPSNLAASGGLAVNSLSPPTPTGLTVASENTIEGTL
GIRGQLPRIQAVNTDGAFTGTGIGAVLYCGCGDAIGIVSEAPIVAPASIGYCOMPVNA
GKRGIPCGGGGLY"

sig_peptide join(757..804,2208..2219)
gene /gene="5H4"
join(757..804,2208..2219)
/gene="5H4"
intron 805..2207
/gene="5H4"
2208..2768
/gene="5H4"
/partial
/gene="5H4"
/number=2
mat_peptide 2220..2681
/gene="5H4"
/product="chorion protein"
BASE COUNT 991 a 645 c 657 g 1024 t
ORIGIN 1061 bp upstream of KpnI site.

Query Match 100.0%; Score 9; DB 21; Length 3317;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2438 TTCGAGAA 2446
|||||||
Cy 25 ttctcagaa 33

RESULT 181
LOCUS VIBAGAB 3323 bp DNA BCT 05-FEB-1999
DEFINITION Vibrio sp. agab gene for beta-agarase, complete cds.
ACCESSION D21202

NID 9498302 GI:498302
VERSION D21202.1
KEYWORDS beta-agarase
SOURCE Vibrio sp. (strain:JT0107) DNA.
ORGANISM Vibrio sp.
REFERENCE 1 (bases 1 to 3323)
AUTHORS Sugano, Y.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1993) to the DDBJ/EMBL/GenBank databases. Yasushi Sugano, Japan Tabacco Inc., Seawater Science Research Laboratory; 4-13-20 Sakawa, Odawara, Kanagawa 256, Japan (Tel:0465-47-3161, Fax:0465-48-6242)
REFERENCE 2 (sites)
AUTHORS Sugano, Y., Matsumoto, T., Kodama, H. and Noma, M.
TITLE Cloning and sequencing of agab, a unique agarase 0107 gene from a marine bacterium, Vibrio sp. strain JT0107
JOURNAL Appl. Environ. Microbiol. 59 (11), 3750-3756 (1993)
MEDLINE 94113702
AUTHORS Sugano, Y., Matsumoto, T. and Noma, M.
TITLE Sequence analysis of the agab gene encoding a new beta-agarase from Vibrio sp. strain JT0107
JOURNAL Biochim. Biophys. Acta 1218 (1), 105-108 (1994)
MEDLINE 94250684
COMMENT Submitted (15-Oct-1993) to DDBJ by:
Yasushi Sugano
Seawater Science Research Laboratory
Japan Tabacco Inc.
4-13-20 Sakawa
Odawara
Kanagawa 256
Japan
Phone: 0465-47-3161
Fax: 0465-48-6242.
Location/Qualifiers
1. .3323
/organism="Vibrio sp."
/strain="JT0107"
/db_xref="taxon:678"
259. .3126
/gene="agab"
259. .3126
/gene="agab"
/EC_number="3.2.1.81"
/codon_start=1
/transl_table=1
/evidence-experimental
/product="beta-agarase"
/protein_id="BA04744.1"
/db_xref="PID:g531276"
/db_xref="PID:g531270"
/db_xref="GI:531270"
/translation="MVEVMKFTKKNKIALSLTLLGYGCGSTPSSDAAGAEVVG
TIDPFESAFAFKKVKDKRAEVSDGVTSGLKVNDSVSEANKFYWPVKVH
PDSCFMNNAKGSLSLIDTNPDPANILKLADNVGSGDGLQLYANVPAGEV
PVEMLENGTKRKLGGYGEKINLRNIVEROIFVGGMDQTVINDPNDVADGTI
EASGEYKVSGRPIPTVASTISFDGQPTPAFDSAAATYELKTDGGLAYKLAT
NATPNTIFKAPQDMSEYGFSLAFLESADPELDLEFRVDAENENGGANGV
DSMSVYTLAPGDGDTYLPFGQTSQIVSGMRAPEPKSKSNAAVSIYNGEGLDS
NIVSFOLYONPTDAEFNIKSVRLINIDADARVEGLIDVOYGFSGEMPKITED
EELETMGLAKMKSLSKTSOMGSRISYGMADGPKLGTGFEETKVDGKMSLVDPGN
LFPATGVDNI RMDQTVITIGHDFADKDRSGKEVASVRSMETWLPEDDVLAEEND
YANNVHSGALKKGVSFSGANLDRKYGTFSEKRYMKDITIDRMVDMGFTLGNMA
DPMFYDKKVAIVANGMIFGDHARISTGNDIWGPIHDPPEPTNYSKANTKKIMTV
DKNDPMWGVFVDEISWGTNRKANDANRYGLVNNLSYDMKSPAKAFTSHLEKYVA
IEDLNTSMGVKVAWSAEFKSFDRSLSKMKKDYAEMLSAKYFSTVRELKYA
LPNHLVGAPEADMGVTPAELAGAPVYDWSYLLYAEADNSKGMKSLKELDKPSII
GEFHGSGTSLGFGIVSAASODRAKKTNTYNSIADNPYFGAMWFOYIDSPITG
RANGENYNTGVFVITDTPYVPIVEAKKKNODYMLRYKK"

BASE COUNT 989 a 714 c 797 g 823 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 3323;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2047 TTCACAGA 2055
|||||
Cp 33 ttccacaga 25

RESULT 182
LOCUS 144522 3348 bp DNA PAT 23-JUL-1997
DEFINITION Sequence 34 from patent US 5635177.
ACCESSION I44522
NID 92469235
VERSION I44522.1 GI:2469235
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3348)
AUTHORS Bennett, B.D., Goeddel, D. and Matthews, W.
TITLE Protein tyrosine kinase agonist antibodies
JOURNAL Patent: US 5635177-A 34 03-JUN-1997;
FEATURES
source 1. .3348
/organism="unknown"
BASE COUNT 924 a 769 c 855 g 800 t
ORIGIN

Query Match 100.0%; Score 9; DB 25; Length 3348;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3183 TTCACAGA 3191
|||||
Cp 33 ttccacaga 25

RESULT 183
LOCUS D49837 3493 bp mRNA VRT 10-FEB-1999
DEFINITION Xenopus laevis mRNA for guanylyl cyclase C, complete cds.
ACCESSION D49837
NID 91850773
VERSION D49837.1 GI:1850773
KEYWORDS guanylyl cyclase C.
SOURCE Xenopus laevis adult intestine cDNA to mRNA, clone:xycc2.1.
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 3493)
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia;
TITLE Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
JOURNAL MacFarland, R.T.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1995) to the DDBJ/EMBL/GenBank databases. Ron T MacFarland, Hokkaido University, Biological Sciences; Sapporo, Hokkaido 060, Japan (E-mail:norio.suzuki@es1.hines.hokudai.ac.jp, Tel:011-706-4908, Fax:011-746-1512)
REFERENCE 2 (bases 1 to 3493)
AUTHORS MacFarland, R.T.
TITLE Unpublished (1995)
JOURNAL Submitted (24-Mar-1995) to DDBJ by:
COMMENT MacFarland, R. Ron
Hokkaido University
Biological Sciences
Sapporo, Hokkaido 060
Japan
Phone: 011-706-4908
Fax: 011-746-1512
Email: norio.suzuki@es1.hines.hokudai.ac.jp.
Location/Qualifiers
1. .3493
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="xycc2.1"

5'UTR
CDS

```

/deb_stage="adult"
/tissue_type="intestine"
1. .38
39. .3236
/EC_number="4.6.1.2"
/codon_start=1
/product="guanylyl cyclase C"
/protein_id="BA00638.1"
/db_xref="PID:d1009251"
/db_xref="PID:g1850774"
/db_xref="GI:1850774"
/translation="MRKHTWCLQIILMTDLLEANCMSGSLTMVIMLNDSTENW
IKAVQAVSIGMHVTVKDLREGIKYINADPQTFNTDLYATPGCVSSGEGVEKLN
LRTRRGCVILGPTCTVATAYOMLSLKNTEGVPILISAGSEGLSCDHRSLARMLLPA
KITFFRQWQYEDFIKPKKQWQYIYKMDGNTSCQFWINALESGVSENNALRKE
ILTBELKAVILOENNHKSNTIMCTSPNDIMLHKVAILPOKXVILILDIPNTVYVD
MSSPYMEVAVLVTPRPSMKISNOTGIKALDENYAGTIDGVLFGHILKFLG
SVDIQNTFSFIDQFNMISIIIGALPLILDAAGRELNLVLVSTANNVTELIQFDT
STNOTVMDISPMFIRKNNRPLSDVQSGPHILITLIVAVIIVLIIIGLVPFRY
RLNELRQKWSHISAELILPMTETSHVSLKIDDKRDSGPRQIKRDKVYL
KDLKDENFTEKQKMLNKLQIDYINLTFGTGYLDNNIYAVIECDKGSRLVDLN
DNISYDGTGFMDEFKISVWYDIAKGMSTYASKCEVHGHLKSTNCVVDGMRVYKTD
FRKSLISPEKDMTAPBHLHOGFSQKGVSYCIATOEIILROETFTTEQCDTKE
KISRVONNGECFPRDLNDSANEREIEYIVLKSCEWDEPRRDEPKIENTLSKI
FSNFHQTESYMDTLIRLQLYSKNLEHVEERTQLYRAERADRLVYLILPGFVY
MSLKEGRVPEFEFEVITYSDIVGFTTICKSAPMEVMDMLNKKNFHLLDHD
VVKVETIGDAYVYVSGLPNNGNNAVDISRMALDICEFMSFELRHLGLPWIMYI
IHSGPCAGVYGIKMPRYCLFGDTVTSRMESTGLPLRIHVKSTIILRRDCEFO
YEVGRGYLKGKPEITTYMTGTETOKNINPTPRPVENORLOSDRAEMIMESLOKE
ADGDKKMPRTVAVSYRNGTLEYLQIGTSDOSTIF"
3'UTR
polya_site 3493
BASE COUNT 1105 a 675 c 752 g 961 t
ORIGIN
3493
3237.3493
Query Match 100.0%; Score 9; DB 24; Length 3493;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2926 TTCTGAGAA 2934
|||||
OY 25 ttctgagaa 33

```

RESULT 184
LOCUS AF012465 3517 bp mRNA VRT 26-AUG-1998
DEFINITION Pleuronectes americanus aminopeptidase N (ampN) mRNA, complete cds.
ACCESSION AF012465
NID 93452125
VERSION AF012465.1 GI:3452125
SOURCE Pleuronectes americanus.
ORGANISM Pleuronectes americanus.
Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
Pleuronectiformes; Pleuronectoidae; Pleuronectidae; Pleuronectes.
1 (bases 1 to 3517)
AUTHORS Douglas, S.E., Gallant, J.W. and Bullerwell, C.E.
TITLE Molecular Investigation of Aminopeptidase N Expression in the
Winter Flounder, *Pleuronectes americanus*
J. Appl. Ichthyol. (1998) In press
2 (bases 1 to 3517)
REFERENCE Douglas, S.E. and Gallant, J.W.
AUTHORS Direct Submission
JOURNAL Submitted (07-JUL-1997) Marine Biology, Institute for Marine
Biosciences, 1411 Oxford Street, Halifax, NS B0J 1G0, Canada
FEATURES
Source
1. .3517
/organism="Pleuronectes americanus"
/db_xref="taxon:8263"
/tissue_type="intestine"
1. .3517
/gene="ampN"

CDS

```

24. .2948
/gene="ampN"
/EC_number="3.4.11.2"
/note="microsomal aminopeptidase; membrane alanyl  
aminopeptidase"
/codon_start=1
/product="aminopeptidase N"
/protein_id="AAC32754.1"
/db_xref="PID:g3452126"
/db_xref="GI:3452126"
/translation="MGKGFYSKANGVGVILGAGALATITIALSVYSQEKAKNNNE  
VSTVGPVGPFGPTTSKPAVTPAPSEAMDKRLKSLVPEHYNVKLMPLRTADP  
ITGLVITFGESTVEKCTEDLDLIHSNKLNYELDLGQMARSAVNSGVKAPAKT  
SLQAVTVQVLYQLDGKLLKQWYLFDFGELADLGGCYRSYVMEGKRAVAT  
OMQPDARAPCEPEPMKATENTTLHDEFTVLSAGQRESQPVIDSKNLOTD  
PEOTKMTYLLAFLVSEFTSINNTVDVILIRFARKAIDAGAVYLSKTPILKF  
FEGYINSSYPLPKSQIALLPDMNAGMNNKGLITREALLXDEYVSSNKKERATI  
IAHELAAHMFGLVTLRWNDLWNEGFASVIEYIGADVAEPDMIKOLIVNDVHRV  
FAVDLAASHPLSAKEDIQREPOISELDAISYKSGAVSRMLSDFLTEDFVWGLR  
TYLKEFAAGNAVYTDLMNHLQMAVATKTLPGSVQDINMTVWLQMGFPVVTINTSG  
EVSOKHFLDPDSEVTAPSPFNKVIPIKMTKRTAOPPYWLQKSAIINNEKTTGV  
DMVLANLVGVGYRVYNDSDNMKLNALSTNHOLIOVINRAQLVDAFENLARKIIP  
TYRALSTKYLINERDVMQSAICNLNLFYLMRDSRYVGPMDVYLEKOVPLDDY  
KTLVDKTVKPLGHDOYNOVNAISQAKTIGHECCLTYKHFKNMTKINPIHPL  
RTTVYCNALVAGAKEMQFANSEFONATLASAEKRLSALACTTQPWLLQRLTLD  
PKIRKQATSTIVYIANNVVGOSLAMPFVDRMSYIFNOYGGGSFNSNLNGVTKR  
KFEFEELQLOQFADNSEVFGSGSLAVDSIERTIGNMKWIAENKONVNLWFLAET  
R"
3'UTR
2949.3517
/gene="ampN"
BASE COUNT 905 a 936 c 876 g 800 t
ORIGIN
936
Query Match 100.0%; Score 9; DB 24; Length 3517;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1492 TTCTGAGAA 1500
|||||
OY 25 ttctgagaa 33

```

RESULT 185
LOCUS E13668 3528 bp DNA PAT 24-JUN-1998
DEFINITION cDNA encoding DNA-binding protein which is induced by jasmonate.
ACCESSION E13668
NID 93252445
VERSION E13668.1 GI:3252445
KEYWORDS JP 1997224672-A/1.
SOURCE Glycine max.
ORGANISM Glycine max.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Rosidae; Fabales; Fabaceae; Papilionaceae; Glycine.
1 (bases 1 to 3528)
AUTHORS Shibata, D., Kato, T. and Ota, H.
TITLE DNA CODING NEW DNA-CONNECTED PROTEIN
JOURNAL Patent: JP 1997224672-A 1 02-SEP-1997;
MITSUI GYOSAI SHOKUBUTSU BIO KENKUSHO: KK
OS Glycine max (soybean)
PN JP 1997224672-A/1
PD 02-SEP-1997
PF 21-FEB-1996 JP 1996033973
PI SHIBATA DAISUKE, KATO TOMOHIKO, OTA HIROYUKI
PC C12N15/09, A01H5/00, C07H21/04, C07K14/415//C12N15/10, C12Q1/68; CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FT source
1. .3528
/organism="Glycine max"

FT /strain='SB-P'
FT 5'UTR 1..449
FT CDS 450..3383
FT /product='DNA-binding protein' FT
FT /gene='SUIP-2'
FT /clone='T6-1-20'
FT 3'UTR 3384..3528
FEATURES
Source 1..3528
/organism='Glycine max'
/db_xref='taxon:3847'
BASE COUNT 875 a 716 c 865 g 1072 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3528;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1373 TTCTCAGAA 1381
|||||||
Cp 33 ttctcagaa 25
RESULT 186
LOCUS 108049 3552 bp PAT 14-NOV-1994
DEFINITION Sequence 5 from Patent EP 0265293.
ACCESSION 108049
NID 9589238
VERSION 108049.1 GI:589238
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3552)
AUTHORS Tabot, S. and Richardson, C.C.
TITLE T7 DNA polymerase
JOURNAL Patent: EP 0265293-A2 5 27-APR-1988;
FEATURES
Source 1..3552
/organism='unknown'
BASE COUNT 905 a 879 c 894 g 874 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3552;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3052 TTCTGAGAA 3060
|||||||
Cp 25 ttctgagaa 33
RESULT 187
LOCUS CHNNAKATPP 3643 bp DNA VRT 21-APR-1997
DEFINITION Gallus gallus Na+/K+-ATPase alpha-1 subunit gene, promoter region.
ACCESSION L43603
NID 9950310
VERSION L43603.1 GI:950310
KEYWORDS ATPase; Na+,K+-ATPase alpha subunit; Na, K-ATPase alpha-1 subunit.
SOURCE Chicken.
ORGANISM Gallus gallus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE Yu, H.-Y., Nettikadan, S., Fambrough, D.M. and Takeyasu, K.
AUTHORS 1 (bases 1 to 3643)
TITLE Negative transcriptional regulation of the chicken Na+/K(+)-ATPase
JOURNAL Biochim. Biophys. Acta 1309 (3), 239-252 (1996)
FEATURES
Source 1..3643
Location/Qualifiers
/organism='Gallus gallus'

/strain='Leghorn'
/sub_species='domesticus'
/db_xref='taxon:9031'
/clone='G-alpha4'
/dev_stage='adult'
/tissue_type='kidney'
prim_transcript 3453..35643
BASE COUNT 897 a 914 c 960 g 872 t
ORIGIN
Query Match 100.0%; Score 9; DB 24; Length 3643;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1142 TTCTCAGAA 1150
|||||||
Cp 33 ttctcagaa 25
RESULT 188
LOCUS 166488 3656 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 8 from patent US 5670367.
ACCESSION 166488
NID 92724465
VERSION 166488.1 GI:2724465
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3656)
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 8 23-SEP-1997;
FEATURES
Source 1..3656
/organism='unknown'
BASE COUNT 1051 a 772 c 869 g 964 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3656;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2140 TTCTCAGAA 2148
|||||||
Cp 33 ttctcagaa 25
RESULT 189
LOCUS HGRMSG 3681 bp DNA BCT 07-JAN-1992
DEFINITION H. giganteus hgDIR & hgIDIM genes and ORF(meth) for restriction
enzyme HgDIR and methyltransferase HgIDIM.
ACCESSION X55140
NID 943480
VERSION X55140.1 GI:43480
KEYWORDS hgIDIM gene; hgDIR gene; methyltransferase; restriction enzyme.
SOURCE Herpetosiphon aurantiacus.
ORGANISM Herpetosiphon aurantiacus
Bacteria; Green non-sulfur bacteria; Chloroflexaceae group;
Herpetosiphon.
REFERENCE 1 (bases 1 to 3681)
AUTHORS Kroege, M.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1990) Kroege M., Inst. Mikro- and
Molekularbiologie, JLU Giessen, Frankfurter Strasse 107, D-W-6300
Giessen, Germany
REFERENCE 2 (bases 1 to 3681)
AUTHORS Dusterhoft, A., Erdmann, D. and Kroege, M.
TITLE Stepwise cloning and molecular characterization of the HgDIR
restriction-modification system from Herpetosiphon giganteus Hpa2
JOURNAL Nucleic Acids Res. 19 (5), 1049-1056 (1991)
FEATURES
MEDLINE 9121283
COMMENT See also X55137-X55143.

FEATURES
SOURCE
1. .3681
Location/Qualifiers
/organism="Herpetosiphon aurantiacus"
/strain="Hpa2"
/db_xref="taxon:65"
676. .1755
/gene="hgIDIR"
676. .1755
/gene="hgIDIR"
/codon_start=1
/transl_table=1
/product="restriction enzyme"
/protein_id="CAA38937.1"
/db_xref="PID:943481"
/db_xref="GI:43481"
/db_xref="SWISS-PROT:P24599"
/translation="MEQLNQPLGKITNSNDFTKKANGKQFNNAFIALACMEF
SONIKPIYILEKRNIEHYFAVDQVQIPLDQAFEFHSHYHPTTELIGKTPAI
DVVISLONSQIINAFELIKLAIPTDNTAIPLNLOCEIIVIRPIVLAIAKVF
OONPLALDILDPICARIGMEDATSIQPIPLCELLYTFIDRYQVQVQIPLQPIF
KTQKSLIHENCIDLPMNSFALAKYFELASIKPSKSTTRERTTWTIKLYDPA
ONGKIDKRTLDRTTENKNDKAPASGAMTKRYMNSPELQNRIRKRSIKHITING
ORYLSPERRDLSAIVSTPGLFEETL"
1752. .2681
/gene="hgIDIM"
1752. .2681
/gene="hgIDIM"
/codon_start=1
/transl_table=1
/product="methyltransferase"
/protein_id="CAA38938.1"
/db_xref="PID:943482"
/db_xref="GI:43482"
/db_xref="SWISS-PROT:P24600"
/translation="MKTIDLPAGCGGMSLFPMQAGFPIYAAVDNMRPAITYQONT
PIHDLAQIDAAVSLIKTHSPCLIGPPCQDSSGKRDGSLGAVLADLCVQPKRTFVIG
AIQPAWIMENVERARLSKIHOQCSMLGDEGSLAQQVADALCGVQPKRTFVIG
HRGSLADLANVQLAQSLIYRDYFGLSDTDYRRHPTREARAIISVEPSPT
IRGVNRPATYRHPKPDAGVSLARPLTKERSLIQTFPLDERFVQTKSEQOMIGN
AVPNNLAFLATSLQAVLNQPRMQQSLSPSF"
2725. .>3681
/note="ORF homologous to E.coli metB"
/codon_start=1
/transl_table=1
/protein_id="CAA38939.1"
/db_xref="PID:9581261"
/db_xref="GI:581261"
/db_xref="SWISS-PROT:P24601"
/translation="MDYGFATRAIHAGQDESVGTAVIPIYOTSTYAGQGDHRCY
EYSTRDPTRTALQTCALAEAKHAIVFSGASGASTTLMMLKAGDHVLCGDVYGG
TYRLEFQWTEHGLSFEDIDMADPEAVRAIKPTRLIMLEPTNLKAPLAIAIR
VAREHGIMTIVDTEFASPYNORPTTLCADAVLHSTKYLIGSHSDVVGAGIMTSENDELY
EKLFKLNAGAVGPPDCMVLRGVKTLSIRMDDERMLATAQFLTEHPVEKVIY
PGLPSEQHNLAPROMRGFGMSIILKGADEAANAVSKTKL"

BASE COUNT 1012 a 765 c 801 g 1103 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 3681;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1434 TTCTGAGAA 1442
|||||
OY 25 ttctgagaa 33

RESULT 190
LOCUS ARO21440 3699 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 120 from patent US 5789650.
ACCESSION ARO21440
NID 93976055
VERSION ARO21440.1 GI:3976055
KEYWORDS

SOURCE
ORGANISM Unknown.
REFERENCE Unknown.
AUTHORS Unclassified.
1 (bases 1 to 3699)
Lomberg, N. and Kay, R. M.
TITLE Transgenic non-human animals for producing heterologous antibodies
JOURNAL Patent: US 5789650-A 120 04-AUG-1998;
FEATURES Location/Qualifiers
1. .3699
/organism="unknown"
BASE COUNT 882 a 995 c 899 g 923 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3699;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3201 TTCTGAGAA 3209
|||||
OY 25 ttctgagaa 33

RESULT 191
LOCUS DROUMPS 3710 bp DNA INV 18-NOV-1993
DEFINITION Drosophila melanogaster (lambda J2 clone, pBSr-1ECO subclone) UMP
synthase (r-1) gene, complete cds.
ACCESSION L00968
NID 9158793
VERSION L00968.1 GI:158793
KEYWORDS UMP synthase.
SOURCE Drosophila melanogaster (strain Canton S) (library: Maniatis Chiron
4A) embryonic DNA.
ORGANISM Drosophila melanogaster
Eukaryote; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 3710)
AUTHORS Eisenberg, M.T., Gathy, K., Vincent, T. and Rawls, J.
TITLE Molecular cloning of the UMP synthase gene rudimentary-like in
Drosophila melanogaster
Mol. Gen. Genet. 222, 1-8 (1990)
JOURNAL 91042409
MEDLINE 2 (bases 1 to 3710)
REFERENCE Eisenberg, M.T., Kirkpatrick, R. and Rawls, J.
AUTHORS Structure of the rudimentary-like gene and UMP synthase in
Drosophila melanogaster
Gene 124, 263-267 (1993)
JOURNAL 93185933
MEDLINE
FEATURES
SOURCE
Location/Qualifiers
1. .3710
/organism="Drosophila melanogaster"
/strain="Canton S"
/db_xref="taxon:7227"
/db_xref="embryonic"
/dev_stage="embryonic"
/tissue="lb="Maniatis Chiron 4A"
/map="9386-7"
615. .865
/gene="r-1"
/note="Determined by primer extension analysis of
embryonic RNA and sequencing of embryonic cDNA clones."
/citation=[2]
/citation=[1]
/number=1
/evidence="experimental"
join(728. .865, 926. .1442, 2100. .2926)
/note="rudimentary-like"
/gene="r-1"
/allele=""
/db_xref="FlyBase:FBgn0003257"
join(728. .865, 926. .1442, 2100. .2926)
/gene="r-1"
/EC_number="2.4.2.10"
/note="Bifunctional protein containing catalytic domains for..

orotidiphosphoribosyltransferase (EC 2.4.2.10) and orotidylatedecarboxylase (EC 4.1.1.23).; putative"

/citation=[2]
/citation=[1]
/codon_start=1
/product="UMP synthase; r-1 gene product"
/protein_id="AAA29012.1"
/db_xref="pid:9158794"
/db_xref="gi:158794"
translation="WAQNSDKRALALKEFINAFKGEDEFMKVGINSPYEDLRVI
VSLGPQOTVSDLVLEHIKDKQSAKHCVCVYALPRATIVSVQGGTPIMLVRRERAK
AVGKIKVEGIEINAGDTCLIVEDVVTSGSSIIDTVLDGEGIVTDAVVVDFRQGG
VANAKHGVHMSLEFTEFLNLTLEAGRIEKTVEAVAKYIAAVOISDGFVREGDK
VTFPANDLORTKLTFTESRANLAKSAVAKRLEPNIJASROTNICLAADLTHADELIDVA
DKGSPYCLIKTHVDIVEDFSKFTADLQALAKRNFLMEDRKADIGNTVSLQYK
GIYKISSMDVTAHTLPERSITLQGLKAGLGKGGKGEVFLLEKMSASGLIDAKY
KENSRIATEGADVDFAGVVCSSDAFAPELQLTPEVKIDEGVDLQGOYQSPFH
VYKRGADIGVGRILKASSPKQAQYRDLMAAYODRVAK"

866. .925
/gene="r-1"
/note="Determined by sequencing embryonic cDNA clones."
/citation=[2]
/number=1
/evidence=experimental
/gene="r-1"
926. .1442
/note="Determined by sequencing embryonic cDNA clones."
/citation=[2]
/db_xref="FlyBase:FBgn0003257"
/number=2
/evidence=experimental
1443. .2099
/gene="r-1"
/note="Determined by sequencing embryonic cDNA clones."
/citation=[2]
/number=2
/evidence=experimental
2100. .3365
/gene="r-1"
/note="Determined by sequencing embryonic cDNA clones."
/citation=[2]
/number=3
/evidence=experimental
3332. .3337
/gene="r-1"
3355. .3360
/gene="r-1"
3365. .3367
/gene="r-1"
/citation=[2]
/citation=[1]
/evidence=experimental

BASE COUNT 1008 a 875 c 914 g 913 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 3710;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 335 TTCTCAGAA 343
|||||
Cp 33 ttctcagaa 25

RESULT 192
LOCUS XLBCADH 3750 bp mRNA VRT 26-JUN-1995
DEFINITION X.laevis XB-cad mRNA for XB-cadherin.
ACCESSION X78546
NID 9468816
VERSION X78546.1 GI:468816
KEYWORDS CAD gene; cadherin.
SOURCE African clawed frog.

ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia;
Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 3750)
AUTHORS Muller,H.A., Kuhl,M., Finemann,S., Schneider,S., van der
Poel,S.Z., Hansen,P. and Wedlich,D.
TITLE Xenopus cadherins: the maternal pool comprises distinguishable
members of the family
JOURNAL Mech. Dev. 47 (3), 213-223 (1994)
MEDLINE 95151580
REFERENCE 2 (bases 1 to 3750)
AUTHORS Kuehl,M.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1994) M. Kuehl, University of Uim, Department of
Biochemistry, Albert-Einstein-allee 11, Oberer Eselsberg, 89081
Uim, FRG
3 (bases 1 to 3750)
Herzberg,F., Wildermuth,V. and Wedlich,D.
Expression of Xbcad, a novel cadherin, during oogenesis and early
development of Xenopus
JOURNAL Mech. Dev. 35 (1), 33-42 (1991)
MEDLINE 92062581
FEATURES
source
Location/Qualifiers
1. .3750
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/dev_stage="gastrula"
/clone_lib="lambda zap II"
3. .2720
/gene="XB-cad"
3. .2720
/gene="XB-cad"
/codon_start=1
/product="XB-cadherin"
/protein_id="CA55292.1"
/db_xref="PID:9468817"
/db_xref="GI:468817"
/db_xref="SWISS-PROT:P33152"
translation="MRLGRVYKFLPKGVANOKOAHMGSTOKFRYSVWLGILCLLQV
VPSINVDVSGCPGPFSSANITYSVNRRELEKGRKLVNVDCTTRHGLYDVDSKF
RVLDPGLVLRVHKLHLSKDRFTISTWDAGIKHSINISVNRKHSRSGEARSSE
LPVLPEKHTGLRKRRKRWVPIPKVSENGEPKRLVQIKSKERLSEMEYITV
OGADTPPEGIFRIEKGWQVTPRLREEEKVLLSHAVSENGASVEEMETVTV
IDQNDNRKFTOPVFRGSRVGVQPGKRVMSATDDSDISNGVIASILKODE
EIPNLFETNETGVISLIGTGLREKPEPTLVQAADLQAGLTREGAVIEITTA
NDNAPIDPKRYTALVPENEGFVQRLSTDLDPCTAANQAYKIRVNGGEFNT
TPESRNGILITAKGLDEEVKQYIQTIVENAVFVPLPTSTAITVYVEDYNEAP
VFVPVSRVDVSEDLTRGEKIVSLVADDPKQQLQKLSYFGNDPAMWLTINKNGIV
TGNDNLESEYVKNNTYVIMLVTDGVPVGTGTLILHVLIDINDGVPSPRVFT
MCDGNLEPOVLTITDADIPNTYPSVLSGSELTKAEIDSGKSTMRSLPTQLRK
GDYSIYVLADANROLTVNATVCOEGKAIKCOEKLAVGFDLPILVILGIIIL
LILSLIILFLPKRKVYKEPLLPEPDROMIFVYGGEGGEEDODYDLSLHSGILA
RPDIRNDVPTLMSVPHYRPRSPNPEIGFIDENIDADNDPTAPYDLSLVFDYE
GSGSEASLSLSSNSNNEHDIVYLVNDMGFRFKLADMTGGDDDDDEE"

BASE COUNT 1053 a 772 c 873 g 1050 t
ORIGIN

Query Match 100.0%; Score 9; DB 24; Length 3750;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 563 TTCTCAGAA 571
|||||
OY 25 ttctcagaa 33

RESULT 193
LOCUS A42478 3754 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 6 from Patent WO9504151.
ACCESSION A42478
NID 92297927
VERSION A42478.1 GI:2297927
KEYWORDS

```

SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 3754)
AUTHORS     Khan, M.A., Hormeche, C.E., Chatfield, S.N. and Dougan, G.
TITLE       VACCINE COMPOSITIONS EXPRESSION OF RECOMBINANT FUSION PROTEINS IN
            ATTENUATED BACTERIA
JOURNAL     Patent: WO 9504151-A 6 09-FEB-1995;
COMMENT     MEDEVA HOLDINGS B V (NL)
            Other publication PL 313979 960805
            Other publication GB 2295394 960529
            Other publication JP 8503602T 960423
            Other publication CA 2168459 950209
            Other publication NO 950348 950328
            Other publication AU 7235794 950228
            Other publication FI 950396 950130
            Other publication AU 4719393 940303.
FEATURES    source
            1..3754
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  978 a 977 c 860 g 939 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3754;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3254 TTCTGAGAA 3262
|||||
QY 25 ttctgagaa 33

RESULT 194
LOCUS      A37074 3754 bp DNA circular PAT 05-MAR-1997
DEFINITION Sequence 17 from Patent W09403615.
ACCESSION  A37074
NID         G2294248
VERSION     A37074.1 GI:2294248
KEYWORDS
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 3754)
AUTHORS     Khan, M.A., Hormeche, C.E., Villarreal-Ramos, B., Chatfield, S.N. and
            Dougan, G.
TITLE       EXPRESSION OF RECOMBINANT FUSION PROTEINS IN ATTENUATED BACTERIA
JOURNAL     Patent: WO 9403615-A 17 17-FEB-1994;
COMMENT     MEDEVA HOLDINGS B V (NL)
            Other publication CA 2141427 940217
            Other publication AU 4719393 940303
            Other publication NO 950348 950328
            Other publication FI 950396 950130
            Other publication JP 8503602T 960423.
FEATURES    source
            1..3754
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  978 a 977 c 860 g 939 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3754;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3254 TTCTGAGAA 3262
|||||
QY 25 ttctgagaa 33

RESULT 195
LOCUS      A37075 3769 bp DNA circular PAT 05-MAR-1997
DEFINITION Sequence 18 from Patent W09403615.

```

```

ACCESSION  A37075 32294249
NID         G2294249
VERSION     A37075.1 GI:2294249
KEYWORDS
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 3769)
AUTHORS     Khan, M.A., Hormeche, C.E., Villarreal-Ramos, B., Chatfield, S.N. and
            Dougan, G.
TITLE       EXPRESSION OF RECOMBINANT FUSION PROTEINS IN ATTENUATED BACTERIA
JOURNAL     Patent: WO 9403615-A 18 17-FEB-1994;
COMMENT     MEDEVA HOLDINGS B V (NL)
            Other publication CA 2141427 940217
            Other publication AU 4719393 940303
            Other publication NO 950348 950328
            Other publication FI 950396 950130
            Other publication JP 8503602T 960423.
FEATURES    source
            1..3769
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  982 a 980 c 863 g 944 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3769;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3269 TTCTGAGAA 3277
|||||
QY 25 ttctgagaa 33

RESULT 196
LOCUS      SWSR3 3775 bp DNA INV 21-MAY-1998
DEFINITION Schistosoma mansoni heat shock transcription factor, exons 6-10 and
            alternatively spliced products, complete cds.
ACCESSION  AF043421
NID         G2895596
VERSION     AF043421.1 GI:2895596
KEYWORDS
SEGMENT
SOURCE      Schistosoma mansoni.
ORGANISM    Schistosoma mansoni.
            Schistosoma mansoni.
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigoida; Schistosomatidae; Schistosomidae; Schistosoma.
REFERENCE   1 (bases 1 to 3775)
AUTHORS     Lanter, F., Ziv, E., Ram, D. and Schechter, I.
TITLE       Different forms of the mRNA encoding the heat-shock transcription
            factor are expressed during the life cycle of the parasitic
            helminth Schistosoma mansoni
JOURNAL     Eur. J. Biochem. 253 (2), 390-398 (1998)
MEDLINE     98316660
REFERENCE   2 (bases 1 to 3775)
AUTHORS     Lanter, F., Ziv, E., Ram, D. and Schechter, I.
TITLE       Direct Submission
JOURNAL     Submitted (20-JAN-1998) The Weizmann Institute of Science, Rehovot
            76100, Israel
FEATURES    source
            1..3775
            /organism="Schistosoma mansoni"
            /strain="Puerto Rican"
            /db_xref="taxon:6183"
            /clone="HSF-G1"
            /dev_stage="adult worm"
            join(AF043419.1:935..1073, AF043419.1:1107..1215,
            AF043419.1:1249..1667, AF043419.1:1702..1846,
            AF043420.1:1:397..737, 998..1296, 1590..1747, 1787..1843,
            2463..2812, 3232..3580)
            /note="alternatively spliced SWSR3 mRNA"
            /product="heat shock transcription factor"
            join(AF043419.1:935..1073, AF043419.1:1107..1215,

```

MRNA

AF043419.1:1249..1667,AF043419.1:1702..1846,
 AF043420.1:1397..612,AF043420.1:1658..737,998..1296,
 1590..1747,1787..1843,2463..2812,3232..3580)
 /note="alternatively spliced HSF1 mRNA"
 /product="heat shock transcription factor"
 join(AF043419.1:935..1073,AF043419.1:1107..1215,
 AF043419.1:1249..1667,AF043419.1:1702..1846,
 AF043420.1:1397..621,AF043420.1:1658..737,998..1296,
 1590..1747,1787..1843,2463..2812,3232..3580)
 /note="alternatively spliced HSF2 mRNA"
 /product="heat shock transcription factor"
 join(AF043419.1:963..1073,AF043419.1:1107..1215,
 AF043419.1:1249..1667,AF043419.1:1702..1846,
 AF043420.1:1397..612,AF043420.1:1658..737,998..1296,
 1590..1747,1787..1843,2463..2800)
 /function="binds to HSE in promoters of Hsp genes"
 /note="HSF1 protein; alternatively spliced product"
 /codon_start=1
 /product="heat shock transcription factor"
 /protein_id="AAC39028.1"
 /db_xref="PID:g2885599"
 /db_xref="GI:2893599"

CDS

CDS

CDS

OTDLSGSHKRPEDIRNLAKIGNTSAGNITDQISDLSGVHLLPSNLQTPVSG
 NDEQDVDTSLVKNPNVODGGRVYYSKVDNETHPTNEEPVGLNVEDGVTDN
 NSLIELNPPCSRDTPYCLDPSGLSDVTPSKLDGVDESICDILLNGDSOTQEVGN
 NQPSSESTLSKRYKEPRKQKQIONRRCIAPRAESDPQVSTTLODIDSLPWEKIOD
 SGDFNLDVONOTENSRSSNGLIYGENEIIPEPAPLNCDDVIOFLRESEITGKKIK
 SDEPLKTVATEHAHSTONGAIDLTSSSAPPTWKARLSANNRLTISSSPKVRKQVPLI
 VLIRQGITPLSTPDCEELE"
 998..1296
 /number=6
 1590..1747
 /number=7
 1787..1843
 /number=8
 2463..2812
 /number=9
 3232..3580
 /number=10
 BASE COUNT 1139 a 666 c 646 g 1324 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 22; Length 3775;
 Best Local Similarity 100.0%; Pred.No. 8,17e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1268 TTCACAGAA 1276
 |||||
 Cp 33 ttccacaga 25
 RESULT 197
 LOCUS TNLMBGL 3786 bp DNA BCT 01-JUL-1997
 DEFINITION T.neapolitana lama gene and bglb gene.
 ACCESSION Z47974
 NID 91143598
 VERSION Z47974.1 GI:1143598
 KEYWORDS beta-glucosidase; bglb gene; LAMA gene; laminarinase.
 SOURCE Thermotoga neapolitana.
 ORGANISM Bacteria; Thermotogales; Thermotoga.
 1 (bases 1 to 3786)
 Dakhova,O.N., Kurepina,N.E., Zverlov,V.V., Svetlichnyl,V.A. and
 Vellikodvorskaya,G.A.
 Cloning and expression in Escherichia coli of Thermotoga
 neapolitana genes coding for enzymes of carbohydrate substrate
 degradation
 JOURNAL Biochem. Biophys. Res. Commun. 194 (3), 1359-1364 (1993)
 MEDLINE 93356813
 REMARK (sites)
 2 (bases 1 to 3786)
 Zverlov,V.V.
 Direct Submission
 Submitted (24-JUN-1995) Zverlov V. V., Institute of molecular
 genetics, Russian Academy of Sciences, Kurchatov Sq. 46, Moscow,
 Russia, 123182
 3 (bases 1 to 3786)
 Zverlov,V.V., Volkov,I.Y., Vellikodvorskaya,T.V. and Schwarz,W.H.
 Highly thermostable endo-1,3-beta-glucanase (laminarinase) lama
 from Thermotoga neapolitana: nucleotide sequence of the gene and
 characterization of the recombinant gene product
 JOURNAL Microbiology 143 (Pt 5), 1701-1708 (1997)
 97312008
 Location/Qualifiers
 1..3786
 /organism="Thermotoga neapolitana"
 /strain="Z2706-MC248"
 /db_xref="taxon:2337"
 1..901
 /gene="bglb"
 <1..901
 /gene="bglb"
 /CC_number="3.2.1.21"
 /pseudo
 /transl_table=11

/codon_start=2
/product="beta-glucosidase"
/db_xref="PID:e135645"
RBS 905..910
918..2883
/gene="lama"
918..2858
/gene="lama"
CDS
/standard_name="1,3-beta-D-glucan-glucanhydrolase"
/EC_number="3.2.1.39"
/citation=(1)
/codon_start=1
/transl_table=11
/evidence=experimental
/product="lamnarinase"
/protein_id="CA88008.1"
/db_xref="PID:e135646"
/db_xref="PID:g1143599"
/db_xref="GI:1143599"
/db_xref="SPTREMBL:060039"
/translation="MKKLVYVLLFPVFLAONILNNGSFDAPILLAGVDIEPPADG
SINTONNVFTNSNGEGKARVNGVLYVTITNGCHTMSVQIISPIREKHKRY
FFKAKASVORNIKIGTAGKGMAYNPETIDSGKVFELGDMKTYEFYWRKOT
DENAREFOLGKSTGTWIDTWIMDVMDVGTLEVSSENEIYTEEDKVEDQL
VMSQEFDDGVIDPNVNFIEIGNHAKGIPGMNALEYTDKNAFVENCGLVEARKE
QVDEGVGYTSRITTEGKFEIKGKIRAKLPKGGIMPALMMLNGVIEGMP
TCGEIDMELMDGDTPTVLRTHAPGPGSGASIVAHLPDEVPDSESDHSTFMD
EMERYVVDQCLYHVLKSKDLALGLERVDHFPFLINAMGYPGYPDETTPQ
RMVYIDYRVKDMNPETITGEVDCERESQDQGEVYEQINNGTFDEPIVNDAN
NPDEWIMQADGDISGARVSDYVDGAYITIEDSGDTMHIQFNOMWIGLKGYT
TISFRAKADTPRINVKILNDHNPWIMYEQVNTTEMOTFTFTYTHPDADDEVQI
SFEIGKEAPFTTIFDVSVSPQ"
2864..2883
terminator
/gene="lama"

BASE COUNT 1131 a 783 c 1023 g 849 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 3786;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2055 TTCTCAGAA 2063
|||||
Cp 33 ttctcagaa 25

RESULT 198 STRATPEPFA 3820 bp DNA BCT 25-OCT-1993
LOCUS Streptococcus faecalis flfO ATPase (atpe, atpf, atph and atpa)
DEFINITION genes, complete cds.
ACCESSION M64265
M64265
M64265.1 GI:153575
VERSION M64265.1 GI:153575
KEYWORDS ATPase alpha subunit; ATPase beta subunit; ATPase delta subunit;
atpf gene; atph gene;
Streptococcus faecalis (ATCC 9790) DNA.
SOURCE Enterococcus faecalis
Organism Enterococcus faecalis
Eubacteria; Firmicutes; Low G+C gram-positive bacteria;
Enterococcus.
1 (bases 1 to 3820)
REFERENCE 1 Kobayashi, H.
AUTHORS Unpublished (1991)
JOURNAL 2 (sites)
REFERENCE Shihata, C., Ehara, T., Tomura, K., Igarashi, K. and Kobayashi, H.
ATTORNS Gene structure of Enterococcus hirae (Streptococcus faecalis) flfO-ATPase, which functions as a regulator of cytoplasmic pH
TITLE J. Bacteriol. 174, 6117-6124 (1992)
JOURNAL 93015650
MEDLINE Location/Qualifiers
FEATURES 1..3820
source /organism="Enterococcus faecalis"
/db_xref="taxon:1351"

gene
134..349
/gene="atpe"
CDS 134..349
/gene="atpe"
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=11
/product="flfO-ATPase gamma-subunit"
/protein_id="AAA26861.1"
/db_xref="PID:g153576"
/db_xref="GI:153576"
/translation="MNYIAALAIINGAIGAGYNGGOVYSKTIESMAROPEMSGLRT
TMEIGVALVEAVPIGVVIALILVFAV"
559..1083
/gene="atpf"
559..1083
/gene="atpf"
CDS 559..1083
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=11
/product="flfO-ATPase beta subunit"
/protein_id="AAA26862.1"
/db_xref="PID:g153577"
/db_xref="GI:153577"
/translation="MLNOLAIENGPMLGNIIVVSGSFLIMFLKHFANGPISDIL
KREOREINDLSAKRSINSKAKMREDOOLASRDAADIKNAKESGELSRONIL
KREOREVARIKSKAQTDLMLERDTLNSVKDVALSLQIAKILINKELSPMHESLI
NOYIGLSSNSTR"
1070..1612
/gene="atph"
1070..1612
/gene="atph"
CDS 1070..1612
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=11
/product="flfO-ATPase delta subunit"
/protein_id="AAA26863.1"
/db_xref="PID:g153578"
/db_xref="GI:153578"
/translation="MKLDKTYGRRYKALFELADISASAEIYOELLRLQYSEIP
GLGNLSVRLPEPHEKRITMDKLVSGFGICVKNPELVYYSNRSELSFMDIYEHRY
NDYKGLLSVKTAVPLSDEQLKLENNVAKTMDQIVELKQIYDSITIGALIVAHN
RVIDGSIRTOLEKMRNOLNR"
1635..3191
/gene="atpa"
1635..3191
/gene="atpa"
CDS 1635..3191
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=11
/product="flfO-ATPase alpha subunit"
/protein_id="AAA26864.1"
/db_xref="PID:g153579"
/db_xref="GI:153579"

BASE COUNT 1280 a 672 c 786 g 1082 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 3820;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3321 TTCTCAGAA 3329

|||||
33 tttctgagaa 25

RESULT 199
LOCUS E07844 3821 bp DNA PAT 29-SEP-1997
DEFINITION Genomic DNA of human nerve growth factor2(NGF-2/NT-3).
ACCESSION E07844
NID 92175977
VERSION E07844.1 GI:2175977
KEYWORDS JP 1994189770-A/4.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 3821)
AUTHORS Shintani, A., Sasada, R. and Igarashi, K.
TITLE EXPRESSION PROMOTER AND ITS USE
JOURNAL Patent: JP 1994189770-A 4 12-JUL-1994;
TAKEDA CHEM IND LTD
COMMENT OS Homo sapiens (human)
PN JP 1994189770-A/4
PD 12-JUL-1994
PE 31-MAY-1993 JP 1993128601
PR 01-JUN-1992 JP 92P 140706, 27-OCT-1992 JP 92P 288717 PI
SHINTANI ASAE, SASADA REIKO, IGARASHI KOICHI
PC C12N15/16, C12N15/53;
CC strandedness: Double;
CC topology: linear;
CC feature is identified by similarity;
FH Key Location/Qualifiers
FT source 1..3821 /organism='Homo sapiens'
FT CDS 3276..3821 /product='human NGF-2(nerve growth factor2)'.
FEATURES
source 1..3821 /organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 880 a 1041 c 1031 g 869 t
ORIGIN
Query Match 100.0%; Score 9; DB 25; Length 3821;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 833 TTCTGAGAA 841
|||||
OY 25 tttctgagaa 33
RESULT 200
LOCUS ATPHYD 3886 bp DNA PLN 23-AUG-1994
DEFINITION A.thaliana phyd gene.
ACCESSION X76609
NID 9452812
VERSION X76609.1 GI:452812
KEYWORDS phyd gene; phytochrome D.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
Rosidae; Capraresae; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 3886)
AUTHORS Sharrrock, R.A.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1993) R.A. Sharrrock, Montana State University,
Dept of Biology, Bozeman, MT 59717, USA
2 (bases 1 to 3886)
REFERENCE Clack, T., Mathews, S. and Sharrrock, R.A.
TITLE The phytochrome apoprotein family in Arabidopsis is encoded by five
genes: the sequences and expression of PHYD and PHYE

JOURNAL Plant Mol. Biol. 25 (3), 413-427 (1994)
MEDLINE 94325466
FEATURES
source Location/Qualifiers
1..3886 /organism='Arabidopsis thaliana'
/db_xref='taxon:3702'
1..1050 /organism='Arabidopsis thaliana'
/strain='landsberg erecta'
/db_xref='taxon:3702'
/clone='lambdaD6-1'
113..142 /note='upstream open reading frame (URF)'
/codon_start=1
/protein_id='CA54070.1'
/db_xref='PID:9579268'
/db_xref='GI:579268'
/translation='MSIIFELLO'
172..207 /note='upstream open reading frame (URF)'
/codon_start=1
/protein_id='CA54071.1'
/db_xref='PID:9452813'
/db_xref='GI:452813'
/translation='MIEDFLRLSH'
242..3736 /gene='phyd'
/gene='phyd'
242..3736 /gene='phyd'
/codon_start=1
/product='phytochrome D'
/protein_id='CA54072.1'
/db_xref='PID:9452814'
/db_xref='GI:452814'
/db_xref='SWISS-PROT:P42497'
/translation='WVGSGSKTSGGSGAASSGHRHSRHTSAEQAQSSANKALRSQNO
QPONHGGGTESTNKAIOQYIVDARLHAFVEQSGSGSFDYSQSLKAPDYSSVEEQ
ITAYLSRIORGTYOPGFCILAVESEFTIIGYSENNAREMGLMSOSPSIEDSEVL
TIGTDLRSLEFSSSYLLERFAVREITLNPIMIHNNNGKPFYALIRYDVGLID
LEPARTDPAISISIGAVOSOKLAVRAISHSIOSLIKLCIDVYLSVRLNGYDV
MYKFFHEDHEVEVAESKRNDEPYTGLHPATDIPRASRLPQNRYRMVYCYAP
VRVQDDRLTQFICLVSTLRAPGCHQWNTNGSLASLMAVINGNEEDGNGVT
GGNSMRMLMGVWCHNTSARCIPPLRYACEFFMQAGLOLMELQALOVSEKRVAT
MOTLLCMLLEDSFAGIVTORPSIMDLVCKGNAFLVQGXYPGLGVPTDSQINDIYE
MLVANHSDSSTGLSDGAGPRAALGDAVCGMAVACITKRPFLPFERSHTEKEIK
WGAKKHPEDKDCGQRMNPRSSPOTFLLEVYKSRCPQPEMADAHSLOILIRDSFE
SEANDSKAAAGAVQPHGDDVQGMDEICGAVAREMRLIETAVPIFANDIDCING
WNAKIAELTGLSVEDANGKSLVRELITREKREIVDRILSCALKGBGRNVEKLTIG
SELQGRKAMFVVVNAKSSKDYLNIVGCVFVGQDVTGRIYWDKFINIQGKALIHSP
NELIPPIFADENKCCLEWNTAMERKLGWPRSEVYIGKLVREVSYSYCRUKGPALRK
FMVILHNAIGGQDTRKPPPEPFDKGEFIOALLUNKRVSIDGKIIGAFGLQIPSE
LOALLEVORROSEFYSRRKELAYIFQVIRKPLSLGLRFTNSLEDMDLNEQKOLLT
SYSCERKISKTIGMDKSIDGSGFLLERPEFTGNTNAYVSQVMVYVERQOLTR
NIPYEVKSMAYTGQIRQOVLAFFLISVRIYAMBSSEVELHPTLNOMADGSAVR
LEFRMAAGEGVPEPKVQDMFHSSRWTSPEGLISVCRKILKIMGVQVYIREERSY
FLIVIELPVPLMMMPSS'
1051..3886 /organism='Arabidopsis thaliana'
/strain='Columbia'
/db_xref='taxon:3702'
/clone='lambda7151'
BASE COUNT 1038 a 801 c 986 g 1061 t
ORIGIN

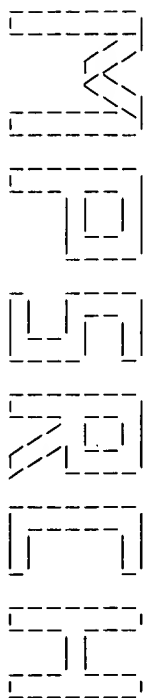
Query Match 100.0%; Score 9; DB 27; Length 3886;
Best Local Similarity 100.0%; Pred. No. 8.17e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1584 TTCTGAGAA 1592
|||||
OY 25 tttctgagaa 33

Thu Oct 28 08:43:22 1999

US-08-963-288-1.rge

Page 80

Search completed: Thu Oct 28 00:25:30 1999
Job time : 369 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mprch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Oct 28 00:30:04 1999; Maspar time 8.94 Seconds
215.813 Million cell updates/sec

Tabular output not generated.

Title: >US-08-963-288-1
Description: (25-33) from US08963288.seq
Perfect Score: 9
N.A. Sequence: 25 ttctgagaa 33
Comp: aagactctt

Scoring table: TABLE default

Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 271905 segs, 107135622 bases x 2

Post-processing: Minimum Match 100%
Listing first 200 summaries

Database:

n-genesegs
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40 41:part41 42:part42 43:part43
44:part44 45:part45 46:part46 47:part47 48:part48
49:part49 50:part50 51:part51 52:part52 53:part53
54:part54 55:part55 56:part56 57:part57 58:part58
59:part59 60:part60

Statistics: Mean 4.232; Variance 1.823; scale 2.321

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description	Pred. No.
C	1	9	100.0	9	29	T41583	Cytokine activated ST	1.37e+03
C	2	9	100.0	9	18	T41419	Cytokine responsive D	1.37e+03
C	3	9	100.0	9	18	43 V24228	Primer alpha(1)b for	1.37e+03
C	4	9	100.0	9	19	49 V56942	Regulatory element co	1.37e+03
C	5	9	100.0	9	19	29 T41612	Oligonucleotide conta	1.37e+03
C	6	9	100.0	9	20	15 G91527	ETS-2 MspI RFLP prime	1.37e+03
C	7	9	100.0	9	20	3	HPV-16 primer (2).	1.37e+03
C	8	9	100.0	9	20	39 T88330	Primer E2-A for murin	1.37e+03
C	9	9	100.0	9	20	59 V80051	Human PBM2 Intron 5/e	1.37e+03

10	9	100.0	22	49	V54907	Primer used to intro	1.37e+03
11	9	100.0	23	32	T72679	Pik-related kinase MC	1.37e+03
12	9	100.0	23	59	X10121	Human diallelic polym	1.37e+03
13	9	100.0	24	32	T59351	Human papilloma virus	1.37e+03
14	9	100.0	26	33	T84431	Interferon-beta gene	1.37e+03
15	9	100.0	26	7	Q47439	Rat G protein, G1-1,	1.37e+03
16	9	100.0	26	7	Q47447	Rat G protein, G1-1,	1.37e+03
17	9	100.0	26	54	V70065	Human SSX1 gene-speci	1.37e+03
18	9	100.0	26	7	Q47445	G protein, G1-1, comm	1.37e+03
19	9	100.0	26	7	Q47451	Human G protein, G1-1	1.37e+03
20	9	100.0	27	30	T43149	HPV type 16 detection	1.37e+03
21	9	100.0	28	24	T30487	Primer 5/kappa24 for	1.37e+03
22	9	100.0	29	30	T43147	HPV type 16 detection	1.37e+03
23	9	100.0	32	32	T59395	Human papilloma virus	1.37e+03
24	9	100.0	32	44	V14136	Dual tag-gal promoter	1.37e+03
25	9	100.0	33	39	V05901	Primer #2 to amplify	1.37e+03
26	9	100.0	33	52	V64521	M. tuberculosis fusio	1.37e+03
27	9	100.0	34	39	V14350	Primer B1052 for gree	1.37e+03
28	9	100.0	34	4	Q27404	Human chromosome 21 c	1.37e+03
29	9	100.0	40	30	T69419	Plasmid p1825f1i comp	1.37e+03
30	9	100.0	41	6	Q37535	Shigella probe 1683.	1.37e+03
31	9	100.0	42	49	V54906	Primer used to intro	1.37e+03
32	9	100.0	42	35	T73035	Mutagenic PCR primer	1.37e+03
33	9	100.0	50	1	N91955	Complementary strand	1.37e+03
34	9	100.0	54	60	V98966	Human EGF-R hairpin r	1.37e+03
35	9	100.0	54	36	T99091	N-terminal primer for	1.37e+03
36	9	100.0	54	54	V64164	Human IL-2 receptor g	1.37e+03
37	9	100.0	60	11	Q62340	Oligonucleotide probe	1.37e+03
38	9	100.0	63	4	Q67744	Epitope YOL1/34 codin	1.37e+03
39	9	100.0	77	2	N60749	Sequence encoding alp	1.37e+03
40	9	100.0	77	32	T79102	RNA ligand to HIV-1 n	1.37e+03
41	9	100.0	79	19	T21579	Human gene signature	1.37e+03
42	9	100.0	80	2	N60751	Sequence encoding gsm	1.37e+03
43	9	100.0	80	9	Q51442	Human FACC Intron 7.3	1.37e+03
44	9	100.0	100	11	Q62342	Oligonucleotide probe	1.37e+03
45	9	100.0	130	11	Q62537	Oligonucleotide probe	1.37e+03
46	9	100.0	149	15	T03931	Human factor-V Intron	1.37e+03
47	9	100.0	195	8	Q61212	Human brain expressed	1.37e+03
48	9	100.0	197	5	Q30026	PCR product prepared	1.37e+03
49	9	100.0	197	5	Q30027	PCR product prepared	1.37e+03
50	9	100.0	198	33	T79775	Potato starch branch	1.37e+03
51	9	100.0	209	20	T72515	Human gene signature	1.37e+03
52	9	100.0	218	10	Q77502	Human genome fragment	1.37e+03
53	9	100.0	243	13	Q73722	Retronosposon probe	1.37e+03
54	9	100.0	251	53	V85500	EST clone HR693.	1.37e+03
55	9	100.0	258	35	T93999	Model pre-therapeutic	1.37e+03
56	9	100.0	261	17	T01037	Mouse B7-2 exon 5.	1.37e+03
57	9	100.0	282	17	Q61834	T-cell receptor beta-	1.37e+03
58	9	100.0	293	52	V69007	DNA molecule encoding	1.37e+03
59	9	100.0	296	54	V87805	EST clone EH61.	1.37e+03
60	9	100.0	297	4	N50452	Hop stunt viroid RNA.	1.37e+03
61	9	100.0	297	40	V16717	Nucleotide sequence	1.37e+03
62	9	100.0	306	10	Q57426	Zinc resistant like p	1.37e+03
63	9	100.0	314	10	Q76420	Human genome fragment	1.37e+03
64	9	100.0	325	5	Q60687	Human brain Expressed	1.37e+03
65	9	100.0	326	54	V87753	EST clone EQ187.	1.37e+03
66	9	100.0	329	20	T21863	Human gene signature	1.37e+03
67	9	100.0	329	8	Q59108	Human brain Expressed	1.37e+03
68	9	100.0	337	23	T26204	Human gene signature	1.37e+03
69	9	100.0	338	43	V24222	Fragment of centromer	1.37e+03
70	9	100.0	339	43	V24223	Fragment of centromer	1.37e+03
71	9	100.0	339	19	T19986	Human gene signature	1.37e+03
72	9	100.0	342	4	Q26400	GBSS gene fragment I	1.37e+03
73	9	100.0	343	8	Q60547	Human brain Expressed	1.37e+03
74	9	100.0	348	52	V69006	DNA molecule encoding	1.37e+03
75	9	100.0	359	53	V88650	EST clone GZ78.	1.37e+03
76	9	100.0	361	8	Q59805	Human brain Expressed	1.37e+03
77	9	100.0	371	37	T89417	Human chemotactic cyt	1.37e+03
78	9	100.0	374	53	V89589	EST clone CN922.	1.37e+03
79	9	100.0	375	45	V56486	Partial nucleotide se	1.37e+03
80	9	100.0	381	30	T67613	H. pylori cytoplasmic	1.37e+03
81	9	100.0	389	54	V87945	EST clone FG965.	1.37e+03
82	9	100.0	412	53	V88795	EST clone H0640.	1.37e+03

83	9	100.0	413.18	T19076	Human gene signature	1.37e+03
84	9	100.0	415.21	T15982	pskRAB fragment #1.	1.37e+03
85	9	100.0	420.48	V44633	Human uncoupling prot	1.37e+03
86	9	100.0	424.31	T64806	Rat RC-9 DNA useful 1	1.37e+03
87	9	100.0	430.38	V02896	Human HMGI-C aberrant	1.37e+03
88	9	100.0	436.52	V34161	Human secreted protein	1.37e+03
89	9	100.0	444.1	N80067	Equine gamma interfer	1.37e+03
90	9	100.0	446.8	O59535	Human brain signature	1.37e+03
91	9	100.0	450.18	T19056	Human gene signature	1.37e+03
92	9	100.0	459.53	V88140	EST clone DQ276.	1.37e+03
93	9	100.0	481.3	O29870	Pharmomone receptor c1	1.37e+03
94	9	100.0	483.51	V63193	cDNA from clone ehb1_	1.37e+03
95	9	100.0	484.53	V88491	EST clone G54.	1.37e+03
96	9	100.0	484.53	V88866	EST clone IA32.	1.37e+03
97	9	100.0	495.11	O63681	Partial cDNA encoding	1.37e+03
98	9	100.0	495.54	V82582	Thyroid hormone recep	1.37e+03
99	9	100.0	498.4	O24293	Ovine gamma-interfero	1.37e+03
100	9	100.0	498.56	V69758	Tobacco ovary-specific	1.37e+03
101	9	100.0	501.3	N50023	Sequence encoding new	1.37e+03
102	9	100.0	501.3	N50027	Sequence encoding new	1.37e+03
103	9	100.0	501.3	N50028	Sequence encoding new	1.37e+03
104	9	100.0	501.3	N50024	Sequence encoding new	1.37e+03
105	9	100.0	501.3	N50025	Sequence encoding new	1.37e+03
106	9	100.0	501.3	N50032	Sequence encoding new	1.37e+03
107	9	100.0	501.3	N50031	Sequence encoding new	1.37e+03
108	9	100.0	501.3	N50023	Sequence encoding new	1.37e+03
109	9	100.0	501.3	N50027	Sequence encoding new	1.37e+03
110	9	100.0	501.3	N50027	Sequence encoding new	1.37e+03
111	9	100.0	501.3	N50025	Sequence encoding new	1.37e+03
112	9	100.0	501.3	N50026	Sequence encoding new	1.37e+03
113	9	100.0	503.20	T27556	Partial human transfo	1.37e+03
114	9	100.0	506.38	V02884	Human HMGI-C aberrant	1.37e+03
115	9	100.0	514.19	T19874	Human gene signature	1.37e+03
116	9	100.0	519.60	X13325	Enterococcus faecalis	1.37e+03
117	9	100.0	521.7	O45770	Human prostate transg	1.37e+03
118	9	100.0	534.54	V87746	EST clone EP234.	1.37e+03
119	9	100.0	539.2	N60632	Beta-urogastrone - be	1.37e+03
120	9	100.0	541.48	T98787	DNA encoding a S. pne	1.37e+03
121	9	100.0	547.52	V68888	DNA molecule encoding	1.37e+03
122	9	100.0	556.3	O22834	Sequence encoding owl	1.37e+03
123	9	100.0	567.60	X13601	Enterococcus faecalis	1.37e+03
124	9	100.0	582.50	V65298	DNA encoding a S. pne	1.37e+03
125	9	100.0	600.54	V88046	EST clone FP268.	1.37e+03
126	9	100.0	603.42	V21154	5' nucleotide portion	1.37e+03
127	9	100.0	613.24	T28030	Mouse incomplete BMP	1.37e+03
128	9	100.0	614.53	V89994	EST clone CM775.	1.37e+03
129	9	100.0	627.16	O09370	Rat allograft inflamm	1.37e+03
130	9	100.0	637.31	T43455	ATM gene exon 15.	1.37e+03
131	9	100.0	666.12	O70205	Influenza haemaggluti	1.37e+03
132	9	100.0	666.12	O70204	Influenza haemaggluti	1.37e+03
133	9	100.0	673.37	T95699	Novel human gene, des	1.37e+03
134	9	100.0	688.1	O05141	Sequence encodes prec	1.37e+03
135	9	100.0	688.11	O67599	Retinoblastoma protei	1.37e+03
136	9	100.0	700.9	O57506	Rat GAP-43 promoter.	1.37e+03
137	9	100.0	720.9	O53372	Sequence encoding glu	1.37e+03
138	9	100.0	746.44	V34780	Human RATH1.1 DNA.	1.37e+03
139	9	100.0	756.22	T14653	Human REST protein DN	1.37e+03
140	9	100.0	776.22	O29389	EG/ET region of Human	1.37e+03
141	9	100.0	790.5	O29389	DNA encoding HPV 16 E	1.37e+03
142	9	100.0	792.43	V17986	Escherichia coli mod1	1.37e+03
143	9	100.0	795.43	V17983	Escherichia coli mod1	1.37e+03
144	9	100.0	813.46	V37407	Streptococcus pneumo	1.37e+03
145	9	100.0	817.46	V40858	Beta-lactamase fragme	1.37e+03
146	9	100.0	827.46	V12222	E. coli J56 pathogeni	1.37e+03
147	9	100.0	832.1	N80115	Human preprolactin ge	1.37e+03
148	9	100.0	843.56	V84414	Human secreted protei	1.37e+03
149	9	100.0	850.8	O48552	Autonomously replicat	1.37e+03
150	9	100.0	852.24	T39913	Maize acetyl CoA carb	1.37e+03
151	9	100.0	858.43	V17984	Escherichia coli mod1	1.37e+03
152	9	100.0	862.51	V59591	Human secreted protei	1.37e+03
153	9	100.0	867.42	V29323	Type A2 ACCase genomi	1.37e+03
154	9	100.0	878.32	T72682	PIK-related kinase MC	1.37e+03
155	9	100.0	890.7	O43705	Sequence of the beta-	1.37e+03

156	9	100.0	892.48	V32990	Human p24 vesicle tra	1.37e+03
157	9	100.0	911.38	V06143	Viral infection gene	1.37e+03
158	9	100.0	918.12	O70208	Sequence encoding imm	1.37e+03
159	9	100.0	924.38	V06123	Viral infection gene	1.37e+03
160	9	100.0	934.38	V06123	Viral infection gene	1.37e+03
161	9	100.0	953.3	N91055	Sequence encoding exo	1.37e+03
162	9	100.0	958.39	T96064	Apoptosis associated	1.37e+03
163	9	100.0	961.14	O79929	Anti-tobacco mosaic v	1.37e+03
164	9	100.0	972.30	T68932	Bovine PIT-1 exon 6	1.37e+03
165	9	100.0	987.45	V59982	Nucleic acid Tc2 enco	1.37e+03
166	9	100.0	993.45	V24131	Homo sapiens Tc152 se	1.37e+03
167	9	100.0	1016.28	T51349	Bacillus subtilis pro	1.37e+03
168	9	100.0	1021.54	V69789	CTLA4-ova fusion prot	1.37e+03
169	9	100.0	1022.18	T13161	-GSTA segment 2.	1.37e+03
170	9	100.0	1026.42	V23014	DNA encoding tryptoph	1.37e+03
171	9	100.0	1039.54	V68168	Porcine interferon-be	1.37e+03
172	9	100.0	1039.54	V08185	Porcine interferon-be	1.37e+03
173	9	100.0	1044.38	V04889	Nucleic acid sequence	1.37e+03
174	9	100.0	1047.37	T88112	Phenylalanyl tRNA syn	1.37e+03
175	9	100.0	1060.10	O57986	Genomic sequence of h	1.37e+03
176	9	100.0	1073.44	V30311	Bacillus thuringiensis	1.37e+03
177	9	100.0	1082.47	T98694	DNA encoding a S. pne	1.37e+03
178	9	100.0	1087.38	T91592	Rat UBC2A coding seq	1.37e+03
179	9	100.0	1097.47	V44860	Clone BG142.1 coding	1.37e+03
180	9	100.0	1100.16	O96139	Prolactin cDNA.	1.37e+03
181	9	100.0	1110.12	O72807	DNA encoding stem reg	1.37e+03
182	9	100.0	1112.3	O21023	mglB-delta-alpha1phal	1.37e+03
183	9	100.0	1119.42	V30002	DNA encoding a Staphy	1.37e+03
184	9	100.0	1129.24	T29786	hmur2 cDNA.	1.37e+03
185	9	100.0	1126.29	T66409	DNA encoding hemangi	1.37e+03
186	9	100.0	1126.29	V08535	DNA encoding hemangi	1.37e+03
187	9	100.0	1151.60	V55785	Mouse B-2 antigen co	1.37e+03
188	9	100.0	1153.14	O81366	Murine B lymphocyte a	1.37e+03
189	9	100.0	1167.47	V37401	Streptococcus pneumo	1.37e+03
190	9	100.0	1176.17	O99950	Recombinant high affi	1.37e+03
191	9	100.0	1180.60	V84163	Streptococcus uberis	1.37e+03
192	9	100.0	1181.15	O85933	Streptococcus uberis	1.37e+03
193	9	100.0	1181.60	V84164	Streptococcus uberis	1.37e+03
194	9	100.0	1185.29	T62556	Human IP-10 gene.	1.37e+03
195	9	100.0	1189.49	V59981	Nucleic acid Tc1 enco	1.37e+03
196	9	100.0	1197.12	O73730	Rape abscission/dehis	1.37e+03
197	9	100.0	1204.44	V37033	Streptococcus pneumo	1.37e+03
198	9	100.0	1212.3	N40167	Sequence encoding gal	1.37e+03
199	9	100.0	1218.49	V18169	Gulf War veteran subj	1.37e+03
200	9	100.0	1221.25	T42978	Human pancreas-deriv	1.37e+03

ALIGNMENTS

RESULT 1
 ID T41583 standard; DNA; 9 BP.
 AC T41583;
 DT 04-JUN-1997 (first entry)
 DE Cytokine activated STAT protein dependent DNA regulatory element.
 KW Regulatory element; protein; cytokine; responsive; host cell;
 KW transfection; agonist; antagonist; mediated; transcription;
 KW modulation; STAT; STAT5; STAT6; signaling pathway; ss.
 OS Synthetic.
 PN WO96030515-A1.
 PD 03-OCT-1996.
 PF 25-MAR-1996; U04012.
 PR 27-MAR-1995; US-411020.
 PA (LIGA-) LIGAND PHARM INC.
 PI Lamb IP, Seidel HM, Tian Chan S;
 DR WPI: 96-455362/45.
 PT DNA construct for screening modulators of cytokine-mediated
 PT transcription - confg. regulatory element and a cytokine-sensitive
 PT promoter operably linked to a heterologous gene
 PT Clam 5; Page 61; 72p; English.
 CC A novel DNA construct comprises an oligonucleotide (ON) comprising
 CC a regulatory element having the present sequence, operably linked
 CC to a promoter, which is operably linked to a heterologous gene
 CC (preferably a marker gene). The gene is under the transcriptional

CC control of the promoter and the ON sequence when the ON is bound by
 CC a STAT protein activated in response to IL-2, IL-3, G-CSF, GM-CSF,
 CC erythropoietin, thrombopoietin, or preferably IL-4, IL-7, IL-9,
 CC IL-13 or IL-15. Cytokine responsive host cells transfected with the
 CC DNA construct can be used to measure the ability of a compound to
 CC act as an agonist or antagonist of cytokine mediated gene
 CC transcription. In particular, they can be used to screen for
 CC cytokine modulators involved in the STAT5 and/or STAT6 protein
 CC signalling pathway.
 CC Sequence 9 BP: 3 A: 2 C: 1 G: 3 T:

Query Match 100.0%; Score 9; DB 29; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 tttcagaa 9
 |||||
 Cp 33 tttcagaa 25

RESULT 2
 ID T14149 standard; DNA: 9 BP.
 AC T14149;
 DT 29-MAY-1996 (first entry)
 DE Cytokine responsive DNA spacer regulatory element.
 KW Regulatory element: transcriptional regulatory protein;
 KW signalling molecule: DNA spacer; agonist: antagonist: anemia;
 KW gene-transcription; inflammation; cytopenia; cancer; ss.
 OS Synthetic.
 PN WO9528482-A2.
 PD 26-OCT-1995.
 PF 10-APR-1995: U04477.
 PR 14-APR-1994: US-228935.
 PR 27-MAR-1995: US-410780.
 PA (LIGA-) LIGAND PHARM INC.
 PI Lamb IP, Seidel HM;
 DR WPI: 95-373797/48.
 PT DNA spacer regulatory elements responsive to cytokine(s) - for
 PT detecting the presence of transcriptional regulatory protein in a
 PT sample
 PS Claim 7; Page 125; 135pp; English.
 CC The present oligonucleotide comprises a regulatory element
 CC TR(Nx)AA, where x is 4-7, and the regulatory element binds an
 CC activated transcriptional regulatory protein in response to a
 CC signalling mol., i.e. a cytokine. This cytokine responsive DNA
 CC spacer regulatory element can be used to detect the presence of
 CC a transcriptional regulatory protein in a sample, and in assays
 CC for (ant)agonists of gene transcription. The identified cpds.
 CC may be used to treat cytokine-induced disease states, or to
 CC ameliorate disease states caused by cytokine deficiency, e.g.,
 CC inflammation, anemia, cytopenia, and (pre)cancerous conditions.
 CC Sequence 9 BP: 3 A: 2 C: 1 G: 3 T:

Query Match 100.0%; Score 9; DB 18; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 tttcagaa 9
 |||||
 Cp 33 tttcagaa 25

RESULT 3
 ID V24228 standard; DNA: 18 BP.
 AC V24228;
 DT 12-AUG-1998 (first entry)
 DE Primer alpha(1)b for human alpha DNA sequence.
 KW CENP-B box; mammalian artificial chromosome; MAC; chromosome 21;
 KW mammalian telomere; centromere; yeast artificial chromosomes; YAC;
 KW somatic gene therapy; PCR primer; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN WO9808964-A1.

PD 05-MAR-1998.
 PF 26-AUG-1996: J02381.
 PR 26-AUG-1996: WO-J02381.
 PA (KNOV) IKENO M.
 PA (MASU/) MASUMOTO H.
 PA (OKAZU/) OKAZAKI T.
 PI Cooke HJ, Grimes BR, Ikono M, Masumoto H, Okazaki T;
 DR WPI: 98-216941/19.
 PT Mammalian artificial chromosomes - prepared using DNA construct
 PT comprising mammalian telomere and centromere containing copies of
 PT CENP-B box sequences
 PS Disclosure: Page 34; 95pp; English.
 CC This sequence is a primer for a human alpha DNA sequence that can be
 CC used in the DNA construct of the invention. The construct is a mammalian
 CC artificial chromosome (MAC), and comprises a mammalian telomere and a
 CC centromere, where the centromere has a DNA sequence containing copies of
 CC the CENP-B box sequence. The invention also relates to methods for
 CC establishing yeast artificial chromosomes (YACs) comprising mammalian
 CC telomeres and centromeres. The method can be used to construct, modify
 CC and stably maintain YACs in yeast cells which have the ability to form
 CC MACs when introduced into mammalian cells. The MACs can replicate
 CC autonomously, be stably maintained extrachromosomally and transmitted
 CC efficiently in mammalian cells. The MACs can be used for basic studies on
 CC organization and function of mammalian chromosomes and also as a vector
 CC to introduce DNA segments (genes) of interest to test their functions in
 CC mammalian cells or bodies, since the genes carried by MACs will neither
 CC be subject to variable expression due to integration position effect nor
 CC cause unpredictable insertion mutation on host chromosomes. Furthermore
 CC MACs will have the capacity to accommodate a DNA segment up to megabases
 CC where an entire large gene or group of genes and regulatory elements
 CC could be included. The MACs can be used for somatic gene therapy or to
 CC generate transgenic mice.
 CC Sequence 18 BP: 3 A: 3 C: 4 G: 8 T:

Query Match 100.0%; Score 9; DB 43; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 tttcagaa 9
 |||||
 Cp 25 tttcagaa 33

RESULT 4
 ID V56942 standard; DNA: 19 BP.
 AC V56942;
 DT 02-DEC-1998 (first entry)
 DE Regulatory element containing oligonucleotide #101.
 KW Cytokine-responsive regulatory; primer; promoter; detection; isolation;
 KW transcriptional control; STAT protein; screening; agonist; ss.
 OS Synthetic.
 PN US5814517-A.
 PD 29-SEP-1998.
 PF 27-MAR-1995: 410779.
 PR 27-MAR-1995: US-410779.
 PR 14-APR-1994: US-228935.
 PA (LIGA-) LIGAND PHARM INC.
 PI Lamb IP, Seidel HM;
 DR WPI: 98-541763/46.
 PT DNA constructs containing cytokine-responsive regulatory elements -
 PT useful in assays for transcription-regulating proteins or gene
 PT transcription agonists or antagonists
 PS Example 1: Column 19-20; 58pp; English.
 CC V56842-V56976 and V61601-V61631 are oligonucleotides used in the
 CC production of constructs comprising a cytokine-responsive regulatory
 CC element linked to a promoter which is linked to a heterologous coding
 CC sequence so that the coding sequence is under the transcriptional control
 CC of the regulatory element and the promoter, where the regulatory element
 CC has a nucleotide sequence selected from TTCNNGAA, TTANYTAA, and TTCNTTAA
 CC where N is A, T, C or G, and Y = 3 or 4. The constructs can be used to
 CC detect or isolate transcription-regulating proteins, e.g. STAT proteins,
 CC in a sample by contacting the sample with the construct so that the
 CC protein binds to the regulatory element, and detecting or separating the

CC resulting complex. The cells can be used in screening assays for agonists
 CC of gene transcription, in which the level of expression of the coding
 CC sequence is measured in the presence and absence of a test compound or
 CC in the presence of the corresponding cytokine.
 SQ Sequence 19 BP; 6 A; 4 C; 5 G; 4 T;

Query Match 100.0%; Score 9; DB 49; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 ttctcagaa 16
 |||||||
 Qy 25 ttctcagaa 33

RESULT 5
 ID T41612 standard; DNA; 19 BP.

AC T41612;
 DT 04-JUN-1997 (first entry)
 DE Oligonucleotide containing core DNA regulatory element.
 KM Regulatory element; STAT; protein; cytokine; responsive;
 KM host cell; transfection; agonist; antagonist; mediated; STAT5;
 KM transcription; modulation; signaling pathway; STAT6;
 KM oligonucleotide; electrophoretic mobility shift assay; EMSA; ds.
 OS Synthetic.

FT Key Location/Qualifiers
 FT misc_feature 5..19
 FT /*tag= a
 FT /note= "core DNA regulatory element"

PN W09630515-A1.
 PD 03-OCT-1996.
 PF 25-MAR-1996; U04012.
 PR 27-MAR-1995; US-411020.
 PA (LIGA-) LIGAND PHARM INC.
 PI Lamb IP, Seidel HM, Tian Chan S;
 DR WPI: 96-45362/45.
 PT DNA construct for screening modulators of cytokine-mediated
 PT transcription - contg. regulatory element and a cytokine-sensitive
 PT promoter operably linked to a heterologous gene
 PS Example 1; Page 26; 728p; English.

CC A novel DNA construct comprises the present oligonucleotide (ON),
 CC which contains a core a regulatory element, operably linked to a
 CC promoter, which is operably linked to a heterologous gene
 CC (preferably a marker gene). The gene is under the transcriptional
 CC control of the promoter and the ON sequence when the ON is bound by
 CC a STAT protein activated in response to IL-2, IL-3, G-CSF, GM-CSF,
 CC erythropoietin, thrombopoietin, or preferably IL-4, IL-7, IL-9,
 CC IL-13 or IL-15. Cytokine responsive host cells transfected with the
 CC DNA construct can be used to measure the ability of a compound to
 CC act as an agonist or antagonist of cytokine mediated gene
 CC transcription. In particular, they can be used to screen for
 CC cytokine modulators involved in the STAT5 and/or STAT6 protein
 CC signalling pathway.
 CC Following an electrophoretic mobility shift assay the DNA construct
 CC was found to bind IL-4 and IL-13 moderately.
 CC Sequence 19 BP; 4 A; 5 C; 4 G; 6 T;

Query Match 100.0%; Score 9; DB 29; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 ttctcagaa 16
 |||||||
 Cp 33 ttctcagaa 25

RESULT 6
 ID Q91527 standard; DNA; 20 BP.

AC Q91527;
 DT 22-DEC-1995 (first entry)
 DE TMS-2 MSP1 RFLP primer.
 KM Restriction fragment length polymorphism; RFLP; point mutation;
 KM mapping; primer; polymerase chain reaction; PCR; ETS-2; ss.

OS Synthetic.
 PN CA2136705-A.

PD 27-MAY-1995.
 PF 25-NOV-1994; 136705.
 PR 26-NOV-1993; US-157269.
 PA (CLAR-) CLARKE INST PSYCHIATRY.
 PI Kennedy JL, Petronis A;
 DR WPI: 95-255407/34.

PT Screening for polymorphism by amplification of pooled nucleic acid
 PT - restriction with endonuclease(s), sepn. of fragments and
 PT comparison of restriction patterns, for detecting disease related
 PT mutation(s), in genetic mapping etc.
 PS Example 5; Page 33; 48pp; English.
 CC The primers given in Q91526-27 were used for the analysis of
 CC the TMS-2 MSP1 PCR-RFLP. Restriction patterns of PCR products
 CC were compared to detect the polymorphism.
 SQ Sequence 20 BP; 6 A; 2 C; 6 G; 6 T;

Query Match 100.0%; Score 9; DB 15; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 ttctcagaa 19
 |||||||
 Qy 25 ttctcagaa 33

RESULT 7
 ID Q23061 standard; DNA; 20 BP.

AC Q23061;
 DT 31-JUL-1992 (first entry)
 DE HPV-16 primer (2).
 KM Primer; probe; PCR; amplification; cervical cancer; condyloma;
 KM Bowen's disease; verruca; oral cavity cancer; pharynx cancer;
 KM laryngeal cancer; detection; ss.

OS Human papilloma virus.
 PN J04058888-A.
 PD 25-FEB-1992.
 PF 28-JUN-1990; 171986.
 PR 28-JUN-1990; JP-171986.
 PA (SHIO) SHIONOGI KK.

DR WPI: 92-111672/14.
 PT Oligonucleotide for the detection of human papilloma virus gene
 PT - and diagnosis of e.g. periviscide, skin cancer, bowel disease,
 PT condyloma, cervix utero cancer, etc.

PS Claim 1; Page 1; 10pp; Japanese.
 CC Oligonucleotides having the sequences in Q23060-67 were prepd.
 CC Amplified fragments ("probes", see Q23068-73) were prepd. by PCR
 CC using each pair of primers. Cloned HPV DNA and various cell strains
 CC were used as the specimens and PCR was performed by using the primers
 CC prepd. above. Confirmation that the amplified band was positive
 CC was by Southern blot hybridisation using the "probes".
 CC The method can be used to detect HPV gene sequences with very high
 CC sensitivity from a minute amt. of sample in a short period.
 CC The objective diseases include cervical cancer, condyloma,
 CC Bowen's disease, verruca, skin cancer, oral cavity cancer, pharynx
 CC cancer, laryngeal cancer and so on.
 SQ Sequence 20 BP; 5 A; 2 C; 8 G; 5 T;

Query Match 100.0%; Score 9; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 ttctcagaa 12
 |||||||
 Qy 25 ttctcagaa 33

RESULT 8
 ID T88330 standard; DNA; 20 BP.

AC T88330;
 DT 22-MAY-1998 (first entry)
 DE Primer E2-A for murine Dp-3 protein E region cDNA.

KW Assaying; cell cycle regulator; E region; DP-3 protein;
 KW nuclear localisation signal; NLS; mouse; murine; PCR primer; ss.
 OS Synthetic.
 OS Mus sp.
 PN WO9743647-A1.
 PD 20-NOV-1997.
 PF 15-MAY-1997; G01324.
 PR 15-MAY-1996; GB-010195.
 PA (MEDI-) MEDICAL RES COUNCIL.
 PI De La Luna S, La Thanque NB.
 DR WPI: 98-009053/01.
 PT Assays for identifying regulators of cell cycle progression -
 PT comprise expressing a protein having a transcription factor nuclear
 PT localisation signal and determining the degree of nuclear
 PT localisation
 PS Example 4: Page 21: 65pp: English.
 CC The present sequence was used in the development of a novel assay
 CC for a putative regulator of cell cycle progression. The assay
 CC comprises expressing in a cell a protein comprising the E region
 CC and sufficient C-terminal residues of a DP-3 protein to provide a
 CC functional nuclear localisation signal (NLS), or the NLS of E2F-1,
 CC and a marker for nuclear localisation. Then the degree of nuclear
 CC localisation in the presence and absence of the putative regulator
 CC is determined. Regulators identified using the assay can be used
 CC as potential regulators of cell proliferation, or as models for
 CC rational drug design. Regulation of the NLS may lead to effects
 CC such as enhanced cell division, blocking of cell cycle progression
 CC or apoptosis. The regulators may also be used to design other
 CC candidate regulatory compounds.
 SQ Sequence 20 BP: 7 A; 3 C; 4 G; 6 T;
 Query Match 100.0%; Score 9; DB 39; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 tctcgagaa 9
 1 tctcgagaa 9
 25 tctcgagaa 33
 RESULT 9
 ID V80051 standard; DNA: 20 BP.
 AC V80051;
 DT 16-MAR-1999 (first entry)
 DE human PMW2 intron 5/exon 6 junction sequence.
 KW Phosphomannomutase-2; PMW2; CDG1; mutation; human; transgenic; assay;
 KW carboxylate-deficient glycoprotein syndrome type 1; drug screening;
 KW Jaeken disease; prenatal diagnosis; ss.
 OS Homo sapiens.
 PN WO9849324-A2.
 PD 05-NOV-1998.
 PF 30-APR-1998; E02593.
 PR 27-JAN-1998; GB-001719.
 PR 30-APR-1997; GB-008851.
 PA (GENZ) GENZYME UK LTD.
 PI Matchis G;
 PI WPI: 99-024063/02.
 DR New DNA encoding human phosphomannomutase or its fragments - used to
 PT detect mutations associated with carboxylate-deficient glycoprotein
 PT syndrome-1, particularly for prenatal diagnosis
 PS Disclosure; Fig 3-6: 104pp: English.
 CC The invention relates to a human phosphomannomutase-2 (PMW2) protein and
 CC the nucleotide sequence encoding the protein. The DNA or its fragments
 CC are used to detect mutation in the PMW2 genes that are associated with
 CC the carboxylate-deficient glycoprotein syndrome type 1 (CDG1). The
 CC sequences can also be used to detect expression of PMW2-related cDNA; to
 CC express PMW2 or its mutants; and to create transgenic animals for use in
 CC drug screening and for studying expression pathways. The expressed
 CC proteins are used to screen for agents that modulate activity of PMW2,
 CC for therapy and to raise specific antibodies (for detecting PMW2 or its
 CC mutants, in competitive or capture assays). Biochemical assays for
 CC phosphomannomutase activity are used to identify possible carriers of CDG1
 CC (Jaeken disease). Measuring enzymatic activity in foetal cells (or in

CC parental leucocytes if such cells are unavailable) and detecting
 CC mutations in the PMW2 gene makes possible a better prenatal diagnosis of
 CC CDG1. Sequences V80046-60 represent PMW2 exon/intron junction sequences.
 SQ Sequence 20 BP: 9 A; 2 C; 2 G; 7 T;
 Query Match 100.0%; Score 9; DB 59; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4 tctcgagaa 12
 1 tctcgagaa 12
 33 tctcgagaa 25
 Cp 33 tctcgagaa 25
 RESULT 10
 ID V54907 standard; DNA: 22 BP.
 AC V54907;
 DT 19-NOV-1998 (first entry)
 DE primer used to introduce E104K mutation into TEM-1 beta-lactamase gene.
 KW TEM-1 beta-lactamase gene; recombination; gene shuffling; PCR primer; ss.
 OS Synthetic.
 PN US5811238-A.
 PD 22-SEP-1998.
 PF 30-NOV-1995; 564955.
 PR 04-MAR-1996; US-537874.
 PR 17-FEB-1994; US-198431.
 PR 30-NOV-1995; US-564955.
 PA (AFRY-) AFRYMAX TECHNOLOGIES NV.
 PI Cramerl A, Stemmer WPC;
 PI WPI: 98-530860/45.
 DR Production of polynucleotides with desired properties - by iterative
 PT selection and recombination
 PS Example 6: Columns 49-50: 74pp: English.
 CC PCR primers V54900-14 were used to introduce mutations into the TEM-1
 CC beta-lactamase gene. The PCR products were used to construct mutant
 CC combinations of the gene, using the method of the invention. The
 CC specification describes a method for evolving a polynucleotide for
 CC acquisition of a desired property. The method comprises providing a
 CC population of variants of the polynucleotide, at least one of which is
 CC in cell-free form, shuffling the variants of the polynucleotide to form
 CC recombinant polynucleotides, selecting or screening for recombinant
 CC polynucleotides that have evolved toward the desired property and
 CC repeating the steps with the selected recombinant polynucleotides until
 CC a recombinant polynucleotide has acquired the desired property. The
 CC recombinant polynucleotide produced by method, or its expression product,
 CC can be formulated as a pharmaceutical.
 SQ Sequence 22 BP: 8 A; 4 C; 3 G; 7 T;
 Query Match 100.0%; Score 9; DB 49; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 13 tctcgagaa 21
 1 tctcgagaa 21
 25 tctcgagaa 33
 Oy 25 tctcgagaa 33
 RESULT 11
 ID T72679 standard; DNA: 23 BP.
 AC T72679;
 DT 16-SEP-1997 (first entry)
 DE PIR-related kinase MCCI1 PCR primer ODH26.
 KW Mammalian cell cycle surveillance 1 beta; MCCI1 beta;
 KW PIR-related kinase; phosphatidylinositol kinase; DNA repair;
 KW ataxia telangiectasia; polymerase chain reaction; PCR; primer; ss.
 OS Synthetic.
 PN WO9718323-A2.
 PD 22-MAY-1997.
 PF 18-NOV-1996; U19337.
 PR 16-NOV-1995; US-558666.
 PR 27-FEB-1996; US-007312.
 PR 21-OCT-1996; US-725304.
 PA (ICOS-) ICOS CORP.

PI Hoekstra MF, Holtzman DA, Keegan KS;
 DR WPI: 97-289299/26.
 PT Phosphatidylinositol kinase-related mammalian cell cycle
 PT surveillance 1 alpha and beta proteins - useful for identifying
 PT modulators and inhibitors for treating, e.g. ataxia telangiectasia
 PS Example 1; Page 25; 21pp; English.
 CC Primers ODH23 (T72674) and ODH26 (T72679) were used to amplify a
 CC portion of the human mammalian cell cycle surveillance 1 (MCCS1)
 CC gene. Primer ODH26 is based on the MCCS1 alpha cDNA sequence (see
 CC also T72664). The primers were used in a PCR-based assay to map
 CC the MCCS1 gene to human chromosome 3. A 237 bp product was
 CC generated from a human/rodent chromosome 3 hybrid. In a second
 CC set of amplification reactions, the same primers were used to
 CC sublocalise the MCCS1 gene to the interval between q21 and q25.1 of
 CC chromosome 3.
 SQ Sequence 23 BP; 5 A; 5 C; 5 G; 8 T;
 Query Match 100.0%; Score 9; DB 32; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 5 ttctcgagaa 13
 QY 25 ttctcgagaa 33
 RESULT 12
 ID X10121 standard; DNA; 23 BP.
 AC X10121.
 DT 24-MAR-1999 (first entry)
 DE Human biallelic polymorphic marker downstream primer #427.
 KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
 KW detection; phenotypic typing; characteristic; infection; hereditary;
 KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
 KW treatment; marker; primer; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN W09820165-A2.
 PD 14-MAY-1998.
 PF 05-NOV-1997: U20313.
 PR 06-NOV-1996: US-030455.
 PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
 PI Hudson T, Lander ES, Wang D;
 DR WPI: 98-286974/25.
 PT New isolated nucleic acid segments from the human genome - used for
 PT determining polymorphic forms for use in e.g. forensics, paternity
 PT testing or phenotypic typing for disease
 PS Claim 16; Page 202; 310pp; English.
 CC X09121-X10268 are allele-specific oligonucleotide primers used in the
 CC isolation of various biallelic polymorphic markers found in the human
 CC genome (represented in X10269-X12937). These primers can be used in a
 CC method for determining polymorphic forms in an individual for use in e.g.
 CC forensics, paternity testing or for phenotypic typing for diseases such
 CC as asgamma/globulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular
 CC dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
 CC hypercholesterolemia, polycystic kidney disease, hereditary
 CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
 CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
 CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
 CC autoimmune diseases, inflammation, cancer, diseases of the nervous
 CC system, infection by pathogenic microorganisms, and characteristics such
 CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
 CC endurance, fertility, and susceptibility or receptivity to particular
 CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
 CC segments can also be used to produce medicaments for the treatment or
 CC prophylaxis of such diseases.
 SQ Sequence 23 BP; 7 A; 4 C; 5 G; 7 T;
 Query Match 100.0%; Score 9; DB 59; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ttctcgagaa 9

CP 33 ttctcgagaa 25
 RESULT 13
 ID T59391 standard; DNA; 24 BP.
 AC T59391;
 DT 20-OCT-1997 (first entry)
 DE Human papilloma virus 16 E7 gene PCR primer.
 KW HPV16; E7 protein; fusion protein; oral carcinoma; cervical cancer;
 KW polymerase chain reaction; immunodetection; ss.
 OS Synthetic.
 OS Synthetic.
 PN DE19627031-A1.
 PD 09-JAN-1997.
 PF 04-JUL-1996: 027031.
 PR 04-JUL-1995: DE-024347.
 PR 04-JUL-1995: DE-024346.
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PI Braspenning J, Frey M, Martens R, Meschede WZ, Pavlita M;
 PI Velhagen I, Zentgraf H;
 DR WPI: 97-066767/07.
 PT Detecting antibodies against human papilloma virus proteins - for
 PT early diagnosis of HPV-mediated carcinoma
 PS Example 1; Page 4; 8pp; German.
 CC Methods for detecting antibodies directed specifically against human
 CC papilloma virus (HPV) proteins are claimed. In particular, the method
 CC is used to detect uncontrolled expression of the HPV E6 and E7 genes
 CC to provide early diagnosis of HPV-associated cancers, especially
 CC cervical and oral carcinoma. A HPV protein is bound to a carrier
 CC and is then incubated with a sample such that any anti-HPV antibodies
 CC become bound. The immobilised antibodies are detected, either
 CC directly or indirectly, by reaction with an anti-idiotypic antibody.
 CC In a specific example, the HPV 16 E7 protein was fused to a C-terminal
 CC sequence of 11 amino acids from the SV40 T antigen for immobilisation.
 CC Part of the cloning procedure involved amplifying the E7 gene using the
 CC PCR primers in T59390 and T59391.
 SQ Sequence 24 BP; 5 A; 4 C; 6 G; 9 T;
 Query Match 100.0%; Score 9; DB 32; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 14 ttctcgagaa 22
 QY 25 ttctcgagaa 33
 RESULT 14
 ID T84431 standard; DNA; 26 BP.
 AC T84431;
 DT 13-NOV-1997 (first entry)
 DE Interferon-beta gene reverse primer for standard PCR.
 KW KIT gene; pig; coat colour; pigmentation; primer; PCR;
 KW polymerase chain reaction; interferon-beta; ss.
 OS Synthetic.
 OS Synthetic.
 PN W09705278-A1.
 PD 13-FEB-1997.
 PF 24-JUL-1996: G01794.
 PR 12-DEC-1995: GB-025364.
 PR 27-JUL-1995: GB-015385.
 PA (DALG-) DALGETY PLC.
 PI Andersson L, Moller MJ, Plastow GS, Sligens KW;
 PI Wales R;
 DR WPI: 97-145712/13.
 PT Determin. of coat colour genotype in pigs by analysis of the KIT gene
 PT - for duplication or deletions, or analysis of KIT protein, used to
 PT establish breeding programmes for pigs of selected colour
 PS Claim 12; Page 39; 49pp; English.
 CC Primers IFN-beta REVERSE (T84431) and IFN-beta CRC FORWARD (T84432)
 CC are based on the pig interferon-beta (IFN-beta) gene. They can be
 CC utilised in an internal standard PCR amplification of pig genomic
 CC DNA, simultaneously with primers (see T84420-27) specific for the
 CC pig KIT gene. DNA amplified from pig genomic DNA by the KIT gene

CC primers is quantitated by comparison to DNA amplified using the CC - IPN-beta primers. The 3 alleles for pig coat colour (I, inhibition of coat colour; I(p), patch; and i, development of colour) can be CC differentiated on the basis of duplication/deletion of the KIT gene. CC This allows breeding of pigs with the desired, usually white, coat CC colour.

SO Sequence 26 BP; 6 A; 5 C; 6 G; 9 T;

Query Match 100.0%; Score 9; DB 33; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 5 ttctcagaa 13
|||||
Oy 25 ttctcagaa 33

RESULT 15
ID 047439 standard; cDNA to mRNA: 26 BP.
AC 047439:
DE 26-JAN-1994 (first entry)
DE Rat G protein, Gi-1, primer RABRGTBP 181.
KW Probe; quantification; human; GTP binding protein; G protein;
KW alpha subunit; specific mRNA; detection; hybridisation; diagnosis;
KW pathophysiology; disease state; hereditary; cancer; infectious;
KW osteodystrophy; pituitary tumour; acromegaly; melanoma cells;
KW diabetes; PCR; ss.
OS Synthetic.
FT nucleotide havl nucleotide having sequence complementary to sequence unique to the mRNA

MO9315221-A.
PN 05-AUG-1993.
PF 29-JAN-1993: U00977.
PR 29-JAN-1992: US-827208.
PR 24-MAR-1992: US-857059.
PR 12-NOV-1992: US-974409.
PA (HITB) HITACHI CHEM CO LTD.
PA (HITB) HITACHI CHEM RES CENT INC.
PI Akiyama T, Cooper A, Mitsuhashi M;
DR WPI: 93-258695/32.
PT Quantitating messenger RNA in sample - using immobilised-poly-
PS Claim 15 and 38; Page 46; 17pp; English.
CC The sequences given in 047433-44 are primers which were used in the
CC quantification of human GTP binding protein (G protein)-specific
CC mRNAs. These primers are derived from human and rat G-protein
CC sequences. These primers were used in conjunction with the method of
CC the invention, in PCR, for the detection and quantification of mRNAs
CC in a sample without the need to purify the mRNA from cells. The
CC claimed method comprises identifying a polynucleotide sequence unique
CC to the mRNA, and immobilising an oligomer complementary to this
CC sequence to an insoluble support. The sample is then incubated with
CC the insoluble support such that the unique sequence will hybridise to
CC the bound oligomer and be immobilised. Non-immobilised components are
CC washed from the support and bound RNA is labelled in such a way that
CC the label is incorporated onto the support relative to the amount of
CC mRNA on the support. The amount of bound label is then determined.
CC This method can be used for the reliable, rapid, simultaneous
CC quantification of multiple varieties of mRNA. It may be used for
CC diagnosing and recognition of pathophysiology of various disease
CC states, eg. hereditary diseases, cancer, and infectious diseases.
CC G proteins are thought to be involved in causing various disease
CC states. A genetic deficiency of Gs protein is the molecular basis of
CC hereditary osteodystrophy. Pituitary tumours in acromegalic patients
CC have been shown to contain mutant Gs proteins. G proteins are also
CC involved in invasive and metastatic melanoma cells, and diabetes.
CC See also 047381-666.

SO Sequence 26 BP; 10 A; 4 C; 6 G; 6 T;

Query Match 100.0%; Score 9; DB 7; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 5 ttctcagaa 13

Cp 33 ttctcagaa 25
|||||

RESULT 16
ID 047447 standard; cDNA to mRNA: 26 BP.

AC 047447:
DE 26-JAN-1994 (first entry)
DE Rat G protein, Gi-1, probe RABRGTBP 226.
KW Probe; quantification; human; GTP binding protein; G protein;
KW alpha subunit; specific mRNA; detection; hybridisation; diagnosis;
KW pathophysiology; disease state; hereditary; cancer; infectious;
KW osteodystrophy; pituitary tumour; acromegaly; melanoma cells;
KW diabetes; PCR; ss.

OS Synthetic.
FT nucleotide havl nucleotide having sequence complementary to sequence unique to the mRNA

MO9315221-A.
PN 05-AUG-1993.
PF 29-JAN-1993: U00977.
PR 29-JAN-1992: US-827208.
PR 24-MAR-1992: US-857059.
PR 12-NOV-1992: US-974409.
PA (HITB) HITACHI CHEM CO LTD.
PA (HITB) HITACHI CHEM RES CENT INC.
PI Akiyama T, Cooper A, Mitsuhashi M;
DR WPI: 93-258695/32.
PT Quantitating messenger RNA in sample - using immobilised-poly-
PS Claim 14 and 38; Page 47; 17pp; English.
CC The sequences given in 047445-51 are probes which were used in the
CC quantification of human GTP binding protein (G protein)-specific
CC mRNAs. These probes are derived from human and rat G-protein
CC sequences. These probes were used in the method of the invention,
CC for the detection and quantification of mRNAs in a sample without
CC the need to purify the mRNA from cells. The claimed method comprises
CC identifying a polynucleotide sequence unique to the mRNA, and
CC immobilising an oligomer complementary to this sequence to an
CC insoluble support. The sample is then incubated with the insoluble
CC support such that the unique sequence will hybridise to the bound
CC oligomer and be immobilised. Non-immobilised components are washed
CC from the support and bound RNA is labelled in such a way that the
CC label is incorporated onto the support relative to the amount of
CC mRNA on the support. The amount of bound label is then determined.
CC This method can be used for the reliable, rapid, simultaneous
CC quantification of multiple varieties of mRNA. It may be used for
CC diagnosing and recognition of pathophysiology of various disease
CC states, eg. hereditary diseases, cancer, and infectious diseases.
CC G proteins are thought to be involved in causing various disease
CC states. A genetic deficiency of Gs protein is the molecular basis of
CC hereditary osteodystrophy. Pituitary tumours in acromegalic patients
CC have been shown to contain mutant Gs proteins. G proteins are also
CC involved in invasive and metastatic melanoma cells, and diabetes.
CC See also 047381-666.

SO Sequence 26 BP; 10 A; 4 C; 6 G; 6 T;

Query Match 100.0%; Score 9; DB 7; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 5 ttctcagaa 13
|||||
Cp 33 ttctcagaa 25

RESULT 17
ID V70065 standard; DNA: 26 BP.

AC V70065:
DE 01-MAR-1999 (first entry)
DE Human SSX1 gene-specific PCR primer SSX 1B(3').
KW SSX1; ssx gene; human; tumour associated antigen; cancer; melanoma;
OS PCR; primer; ss.
OS Synthetic.
OS Homo sapiens.

PN M09850528-A1.
 PD 12-NOV-1998.
 PF 25-FEB-1998; U03661.
 PR 05-MAY-1997; US-851138.
 PA (LUDWIG) LUDWIG INST CANCER RES.
 PI Chen Y, Gure AO, Knuth A, Old LJ, Pfeundschnuh M,
 PI Sahin U, Scanlan MJ, Tsang S, Tureci O.
 DR WPI: 98-610379/51.
 PT New SSX gene family members - useful for assaying for cancer cells
 PS Claim 15; Page 8; 19pp; English.
 CC This is the nucleotide sequence of primer SSX 1B(3'), which
 CC corresponds to nucleotide sequence 440-465 of the human SSX1 gene. It is
 CC used with primer SSX 1A(5') (see V70064) in PCR amplifications
 CC in order to determine expression of the SSX1 gene in normal and
 CC tumour tissues. Gene-specific primer pairs (see V70064-73) are
 CC provided for SSX genes 1-5. Determination of SSX gene expression
 CC provides a means for assaying for cancers such as melanoma.
 SQ Sequence 26 BP; 9 A; 5 C; 2 G; 10 T;
 Query Match 100.0%; Score 9; DB 54; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 17 ttctcagaa 25
 CP 33 ttctcagaa 25
 RESULT 18
 ID 047445 standard; cDNA to mRNA; 26 BP.
 AC 047445;
 DT 26-JAN-1994 (first entry)
 DE G protein, G1-1, common probe G11-735.
 KM Probe: quantification: human; GTP binding protein; G protein;
 KM alpha subunit; specific mRNA; detection; hybridisation; diagnosis;
 KM pathophysiology; disease state; hereditary; cancer; infectious;
 KM osteodystrophy; pituitary tumour; acromegaly; melanoma cells;
 KM diabetes; PCR; ss.
 OS Synthetic.
 FT nucleotide hav1 nucleotide having sequence complementary to sequence unique to
 FT the mRNA
 PN M09315221-A.
 PD 05-AUG-1993.
 PF 29-JAN-1993; U00977.
 PR 29-JAN-1992; US-827208.
 PR 24-MAR-1992; US-857059.
 PR 12-NOV-1992; US-974409.
 PA (HITB) HITACHI CHEM CO LTD.
 PA (HITB) HITACHI CHEM RES CENT INC.
 PI Akitaya T, Cooper A, Mitsuhashi M.
 DR WPI: 93-258695/32.
 PT Quantitating messenger RNA in sample - using immobilised-poly-
 PS Example 5; Page 47; 177pp; English.
 CC The sequences given in 047445-51 are probes which were used in the
 CC quantification of human GTP binding protein (G protein)-specific
 CC mRNAs. These probes are derived from human and rat G-protein
 CC sequences. These probes were used in the method of the invention,
 CC for the detection and quantification of mRNAs in a sample
 CC without the need to purify the mRNA from cells. The claimed method
 CC comprises identifying a polynucleotide sequence unique to the mRNA,
 CC and immobilising an oligomer complementary to this sequence to an
 CC insoluble support. The sample is then incubated with the insoluble
 CC support such that the unique sequence will hybridise to the bound
 CC oligomer and be immobilised. Non-immobilised components are washed
 CC from the support and bound RNA is labelled in such a way that the
 CC label is incorporated onto the support relative to the amount of
 CC mRNA on the support. The amount of bound label is then determined.
 CC This method can be used for the reliable, rapid, simultaneous
 CC quantification of multiple varieties of mRNA. It may be used for
 CC diagnosing and recognition of pathophysiology of various disease
 CC states, eg. hereditary diseases, cancer, and infectious disease.
 CC G proteins are thought to be involved in causing various disease
 CC states, eg. hereditary diseases, cancer, and infectious disease.
 CC G proteins are thought to be involved in causing various disease
 CC states. A genetic deficiency of Gs protein is the molecular basis of
 CC states. A genetic deficiency of Gs protein is the molecular basis of

CC hereditary osteodystrophy. Pituitary tumours in acromegalic patients
 CC have been shown to contain mutant Gs proteins. G proteins are also
 CC involved in invasive and metastatic melanoma cells, and diabetes.
 CC See also 047381-666.
 SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T;
 Query Match 100.0%; Score 9; DB 7; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 5 ttctcagaa 13
 CP 33 ttctcagaa 25
 RESULT 19
 ID 047451 standard; cDNA to mRNA; 26 BP.
 AC 047451;
 DT 26-JAN-1994 (first entry)
 DE Human G protein, G1-1, probe HUMGNBPA 234.
 KM Probe: quantification: human; GTP binding protein; G protein;
 KM alpha subunit; specific mRNA; detection; hybridisation; diagnosis;
 KM pathophysiology; disease state; hereditary; cancer; infectious;
 KM osteodystrophy; pituitary tumour; acromegaly; melanoma cells;
 KM diabetes; PCR; ss.
 OS Synthetic.
 FT nucleotide hav1 nucleotide having sequence complementary to sequence unique to
 FT the mRNA
 PN M09315221-A.
 PD 05-AUG-1993.
 PF 29-JAN-1993; U00977.
 PR 29-JAN-1992; US-827208.
 PR 24-MAR-1992; US-857059.
 PR 12-NOV-1992; US-974409.
 PA (HITB) HITACHI CHEM CO LTD.
 PA (HITB) HITACHI CHEM RES CENT INC.
 PI Akitaya T, Cooper A, Mitsuhashi M.
 DR WPI: 93-258695/32.
 PT Quantitating messenger RNA in sample - using immobilised-poly-
 PS Claim 14 and 38; Page 47; 177pp; English.
 CC The sequences given in 047445-51 are probes which were used in the
 CC quantification of human GTP binding protein (G protein)-specific
 CC mRNAs. These probes are derived from human and rat G-protein
 CC sequences. These probes were used in the method of the invention,
 CC for the detection and quantification of mRNAs in a sample without
 CC the need to purify the mRNA from cells. The claimed method comprises
 CC identifying a polynucleotide sequence unique to the mRNA, and
 CC immobilising an oligomer complementary to this sequence to an
 CC insoluble support. The sample is then incubated with the insoluble
 CC support such that the unique sequence will hybridise to the bound
 CC oligomer and be immobilised. Non-immobilised components are washed
 CC from the support and bound RNA is labelled in such a way that the
 CC label is incorporated onto the support relative to the amount of
 CC mRNA on the support. The amount of bound label is then determined.
 CC This method can be used for the reliable, rapid, simultaneous
 CC quantification of multiple varieties of mRNA. It may be used for
 CC diagnosing and recognition of pathophysiology of various disease
 CC states, eg. hereditary diseases, cancer, and infectious disease.
 CC G proteins are thought to be involved in causing various disease
 CC states. A genetic deficiency of Gs protein is the molecular basis of
 CC hereditary osteodystrophy. Pituitary tumours in acromegalic patients
 CC have been shown to contain mutant Gs proteins. G proteins are also
 CC involved in invasive and metastatic melanoma cells, and diabetes.
 CC See also 047381-666.
 SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T;
 Query Match 100.0%; Score 9; DB 7; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 5 ttctcagaa 13
 CP 33 ttctcagaa 25

```

RESULT 20
ID T43149 standard; DNA: 27 BP.
AC T43149;
DE 22-JUL-1997 (first entry)
KW HPV type 16 detection probe 2'.
KW Probe: amplification: detection; Human Papilloma virus; HPV; type 16;
KW oligonucleotide: probe complement pair; OPCR: hybridisation; cancer;
KW non-enzymatic; point mutation; diagnosis; blood typing; identification;
KW tissue classification-HLA-typing; sex determination; paternity testing;
KW forensic science; autoimmune disease; ss.
OS Synthetic.
FH Key
FT misc_feature 1..27
FT /tag- a
FT /note- "Targeting sequence C'"
FT modified_base 1
FT /tag- b
FT /note- "Modified by a chemical functionality group
FT attached to a substituent group of the uridine
FT residue, preferably attached to the C2' position
FT of the ribose moiety uridine, serves as a
FT dienophile in the Diels-Alder reaction between
FT this group and the modification of T43146"
PN M09634984-A1.
PD 07-NOV-1996;
PR 30-APR-1996; U06042.
PR 01-MAY-1995; US-431527.
PA (BIRA ) BIO-RAD LAB INC.
PI Segev D;
DR WPI: 96-506183/50.
PT Amplification and detection of target nucleic acid sequences -
PT comprises chemical linkage of oligo:nucleotide probe complement
PT pairs, esp. useful for detecting single point mutation(s)
PS Example 1: Page 62; 128pp; English.
CC The sequences given in T43146-49 represent probes which were used in the
CC amplification and detection of a 54 bp sequence contained in the Human
CC Papilloma virus (HPV) type 16 genome using the method of the invention.
CC The method uses two oligonucleotide probe complement pairs (OPCPs). Each
CC member of the probe pair contains a chemical functionality group which
CC permits linkage of the probes when the functionality groups are adjacent
CC to one another following hybridisation of the probe pairs to the
CC template. One probe in each OPCP is composed of two regions; the first
CC region hybridises to the target and contains the chemical functionality
CC group while the second region prevents target independent joining. The
CC other probe in the OPCP contains the corresponding chemical functionality
CC group. When two probes in a OPCP are joined, they can serve as template
CC for the second OPCP, which can in turn serve as template for unjoined
CC first probe pairs in a cyclic amplification process. The method provides
CC for the non-enzymatic amplification and detection of target nucleic acid
CC sequences. It can identify minute sequence alterations in genes such as
CC point mutations. The method can be used for the diagnosis of specific
CC disease linked mutations in known gene regions, for blood typing, tissue
CC classification-HLA-typing, sex determin., paternity testing, forensic
CC science or testing of susceptibility to diseases such as autoimmune
CC diseases or cancers. It can also be used for identifying the presence
CC of a foreign genetic sequence e.g. for detecting specific bacterial or
CC viral sequences in plant or animal samples.
SQ Sequence 27 BP: 5 A; 9 C; 5 G; 7 T; 1 U;

Query Match 100.0%; Score 9; DB 30; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DE 11-FEB-1997 (first entry)
DE Primer 5'/kappa24 for HNK-20 V kappa chain coding sequence.
KW Antibody: HNK-20; variable kappa chain; hybridoma; murine; IgA; mouse;
KW F glycoprotein; respiratory syncytial virus; RSV; constant region gene;
KW chimeric antibody; isotype-switched antibody; therapy; infection; human;
KW pneumonia; bronchiolitis; animal; polymerase chain reaction; primer; PCR;
KW amplify; ss.
OS Synthetic.
FH Key
FT misc_feature 1..29
FT /tag- a
FT /note- "Targeting sequence C'"
FT modified_base 29
FT /tag- b
FT /note- "Modified by a chemical functionality group
FT attached to a substituent group of the uridine
FT residue, preferably attached to the C2' position
FT of the ribose moiety uridine, serves as a
FT dienophile in the Diels-Alder reaction between
FT this group and the modification of T43146"
PN M09634984-A1.

Query Match 100.0%; Score 9; DB 24; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


PD 07-NOV-1996.
 PF 30-APR-1996: U06042.
 PR 01-MAY-1995: US-433527.
 PA (BIRA) BIO-RAD LAB INC.
 PI Segev D;
 DR WPI: 96-506183/50.
 PT Amplification and detection of target nucleic acid sequences -
 PT comprises chemical linkage of oligo:nucleotide probe complement
 PT pairs; esp. useful for detecting single point mutation(s)
 PS Example 1: Page 62; 128pp; English.
 CC The sequences given in T43146-49 represent probes which were used in the
 CC amplification and detection of a 54 bp sequence contained in the Human
 CC Papilloma virus (HPV) type 16 genome using the method of the invention.
 CC The method uses two oligonucleotide probe complement pairs (OPCPs). Each
 CC member of the probe pair contains a chemical functionality group which
 CC permits linkage of the probes when the functionality groups are adjacent
 CC to one another following hybridisation of the probe pairs to the
 CC template. One probe in each OPCP is composed of two regions: the first
 CC region hybridises to the target and contains the chemical functionality
 CC group while the second region prevents target independent joining. The
 CC other probe in the OPCP contains the corresponding chemical functionality
 CC group. When two probes in a OPCP are joined, they can serve as template
 CC for the second OPCP, which can in turn serve as template for unjoined
 CC first probe pairs in a cyclic amplification process. The method provides
 CC for the non-enzymatic amplification and detection of target nucleic acid
 CC sequences. It can identify minute sequence alterations in genes such as
 CC point mutations. The method can be used for the diagnosis of specific
 CC disease linked mutations in known gene regions, for blood typing, tissue
 CC classification-HLA-typing, sex determin., paternity testing, forensic
 CC science or testing of susceptibility to diseases such as autoimmune
 CC diseases or cancers. It can also be used for identifying the presence
 CC of a foreign genetic sequence e.g. for detecting specific bacterial or
 CC viral sequences in plant or animal samples.
 SQ Sequence 29 BP; 9 A; 5 C; 9 G; 5 T; 1 U;
 Query Match 100.0%; Score 9; DB 30; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 4 ttctcgagaa 12
 |||||||
 QY 25 ttctcgagaa 33
 RESULT 23
 ID T59395 standard; DNA; 32 BP.
 AC T59395:
 DT 20-OCT-1997 (first entry)
 DE Human papilloma virus 16 E7 gene PCR primer.
 KM HPV16; E7 protein; fusion protein; oral carcinoma; cervical cancer;
 KM polymerase chain reaction; immunodetection; ss.
 OS Synthetic.
 PN DE19627031-A1.
 PD 09-JAN-1997.
 PE 04-JUL-1996: 027031.
 PR 04-JUL-1995: DE-024347.
 PR 04-JUL-1995: DE-024346.
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 PI Braspenning J, Frey M, Martens R, Meschede WZ, Pawlita M;
 PI Velhagen I, Zenggraf H;
 DR WPI: 97-068767/07.
 PT Detecting antibodies against human papilloma virus proteins - for
 PT early diagnosis of HPV-mediated carcinoma
 PS Example 4: Page 6; 8pp; German.
 CC The human papilloma virus (HPV) 16 E7 gene was amplified using the PCR
 CC primers in T59394 and T59395 as part of the procedure for producing E7
 CC protein fused at its N-terminal to 6 histidine residues. The recombinant
 CC fusion protein was expressed in E.coli hosts and was recovered in
 CC denatured form. The HPV E7 protein was subsequently refolded, e.g. by
 CC dissolving in buffers containing urea, then dialysing against buffers of
 CC decreasing urea content.
 SQ Sequence 32 BP; 9 A; 3 C; 10 G; 10 T;

Query Match 100.0%; Score 9; DB 32; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 18 ttctcgagaa 26
 |||||||
 QY 25 ttctcgagaa 33
 RESULT 24
 ID V31436 standard; DNA; 32 BP.
 AC V31436;
 DT 14-SEP-1998 (first entry)
 DE Dual tac-gal promoter 3' PCR primer.
 KM Vector; dual promoter; tac promoter; gale promoter;
 KM UDPgalactose 4-epimerase; plasmid pTCK; PCR; primer; ss.
 OS Synthetic.
 OS Streptococcus thermophilus.
 PN WO9820111-A1.
 PD 14-MAY-1998.
 PF 07-NOV-1997: U20528.
 PR 08-NOV-1996: US-029545.
 PA (CYTE-) CYTEL CORP.
 PI Heranson G, Schultz J;
 DR WPI: 98-286927/25.
 PT Dual bacterial promoter for increasing expression of heterologous
 PT polypeptide - containing components of tac and gal promoters,
 PT especially for producing enzymes for oligosaccharide synthesis
 PS Example 1: Page 16; 44pp; English.
 CC This 3' primer, which contains a XbaI site, was used with a 5'
 CC primer (see V31435) in the PCR amplification of a DNA fragment
 CC containing the tac promoter approximately 1 kb upstream of the
 CC galactose-inducible promoter from the UDPgalactose-4-epimerase
 CC gene (gale) of Streptococcus thermophilus. Plasmid pHP1/tac
 CC was used as template. The 5' primer was designed to hybridise to
 CC vector pBR322 upstream of the promoter on pHP1/tac, while the 3'
 CC primer was designed to hybridise to a sequenced region of the gale
 CC promoter containing the ribosome binding site. The amplified 1.3
 CC kb product, which encompassed both the tac and gale promoters (see
 CC V31431), was inserted into pPHOX2. The invention provides novel
 CC improved expression vectors comprising a dual bacterial promoter
 CC operably linked to a heterologous nucleic acid which encodes a
 CC desired polypeptide. The constructs are useful for expressing the
 CC desired polypeptide in bacterial host cells at high levels.
 SQ Sequence 32 BP; 12 A; 8 C; 2 G; 10 T;
 Query Match 100.0%; Score 9; DB 44; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 15 ttctcgagaa 23
 |||||||
 Cp 33 ttctcgagaa 25
 RESULT 25
 ID V05901 standard; DNA; 33 BP.
 AC V05901;
 DT 01-JUN-1998 (first entry)
 DE Primer #2 to amplify hTNF-alpha-Da construct.
 KM Primer; PCR; amplification; human tumour necrosis factor-alpha; E. coli;
 KM hTNF-alpha; extracellular domain; deletion mutant; expression vector; ss.
 OS Synthetic.
 OS Homo sapiens.
 PN CN1100464-A.
 PD 22-MAR-1995.
 PF 19-MAY-1994: 112172.
 PR 19-MAY-1994: CN-112172.
 PA (UYFU-) UNIV FUDAN.
 PI Cai W, Li C, Zhao S;
 DR WPI: 97-245915/23.
 PT Human tumor necrosis α derivative and its preparing method
 PS Claim 2; Page 2; 11pp; Chinese.

CC Primers V05900-V05901 are used to amplify a construct encoding a
 CC derivative of the human tumour necrosis factor- α (hTNF- α),
 CC designated hTNF- α -Da. The novel derivative contains the sequence
 CC encoding the extracellular domain 4 of TNF- α deleted and replaced
 CC by a synthetic oligonucleotide (V05902). The amplified sequence is
 CC cloned into a vector for expression in *E. coli*.
 SQ Sequence 33 BP; 12 A; 8 C; 7 G; 6 T;

Query Match 100.0%; Score 9; DB 39; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 ttctcagaa 15
 |||||
 CP 33 ttctcagaa 25

RESULT 26
 ID V64521 standard; DNA; 33 BP.
 AC V64521:
 DT 27-JAN-1999 (first entry)
 DE M. tuberculosis fusion protein TbrA3/38kD/Tb38-1 PCR primer #6.
 KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
 KW vaccine; pharmaceutical; infection; diagnosis; PCR primer; ss.
 OS Synthetic.
 OS Mycobacterium tuberculosis.
 PN WO9816646-A2.
 PD 23-APR-1998.
 PF 07-OCT-1997; 018293.
 PR 13-MAR-1997; US-818112.
 PR 11-OCT-1996; US-730510.
 PA (CORI-) CORIXA CORP.
 PI Campos-Neto A, Dillon DC, Houghton R, Iodes MJ,
 PI Reed SG, Skeiky RM, Twardzik DR, Vedrick TS;
 DR WPI: 98-261042/23.
 PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
 PT to develop products for the detection of M. tuberculosis infection
 PT and for diagnosis, treatment and prevention of tuberculosis
 PS Example 7: Page 154; 230pp; English.
 CC V64516-V64521 are PCR primers used in the construction of a fusion
 CC protein containing the immunogenic polypeptides TbrA3, 38kD and Tb38-1
 CC from Mycobacterium tuberculosis (MT). This fusion protein can be used in
 CC a method for inducing protective immunity against tuberculosis (TB). This
 CC sequence can be formulated into vaccines and/or pharmaceutical
 CC compositions for immunising against M. tuberculosis infection or may be
 CC used for the diagnosis of tuberculosis.
 CC Sequence 33 BP; 11 A; 7 C; 9 G; 6 T;

Query Match 100.0%; Score 9; DB 52; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 ttctcagaa 16
 |||||
 CP 33 ttctcagaa 25

RESULT 27
 ID V14350 standard; DNA; 34 BP.
 AC V14350:
 DT 19-MAY-1998 (first entry)
 DE Primer Bio2 for green fluorescent protein coding sequence.
 KW Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish;
 KW genetic engineering marker; gene therapy; PCR primer; amplify; ss.
 OS Synthetic.
 OS Aequorea victoria.
 PN WO9742330-A1.
 PD 13-NOV-1997.
 PF 07-MAY-1997; 007625.
 PR 08-MAY-1996; US-646538.
 PA (USSH) US SEC. DEPT. HEALTH.
 PA (USSH) US DEPT. HEALTH & HUMAN SERVICES.
 PI Galtanaris GA, Pavlakis GN, Stauber RH, Vournakis JN;

DR WPI: 97-558982/51.
 PT New nucleic acid encoding proteins of Aequorea victoria with
 PT increased fluorescence - useful as markers for detecting cellular
 PT transformation, subcellular localisation of proteins, for assessing
 PT gene therapy, mutagenicity etc.
 PS Example 10; Page 56; 105pp; English.
 CC This sequence is a primer for the coding sequence for the wild type green
 CC fluorescent protein (GFP) of Aequorea victoria. The amplified sequence is
 CC mutated to produce the nucleic acid of the invention that encodes a
 CC protein that is (a) a protein that has Leu at position 65 (and optionally
 CC Thr at 168 and optionally further Cys at 66) and has cellular
 CC fluorescence at least 5 times that of wild-type GFP or (b) is a blue
 CC fluorescent protein (BFP) with His at position 67 and also at least one
 CC of Leu at 65 and Ala at 164, and has cellular fluorescence at least 5
 CC times that of BFP (Tyr67 to His). The nucleic acids can be used as
 CC markers in genetic engineering and gene therapy. They may also be used to
 CC detect and characterise regulatory and coding sequence elements that
 CC control subcellular expression and targeting of proteins. Typical
 CC applications are monitoring targeting and transport of proteins in cells;
 CC assessment of gene therapy procedures; in diagnosis (when expressed under
 CC control of a promoter induced by a particular analyte); assessment of
 CC mutagenicity of compounds; and for drug screening (where expression is
 CC controlled by the promoter of a target gene), particularly for antiviral
 CC or antiparasitic agents.
 SQ Sequence 34 BP; 12 A; 8 C; 9 G; 5 T;

Query Match 100.0%; Score 9; DB 39; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 ttctcagaa 15
 |||||
 CP 33 ttctcagaa 25

RESULT 28
 ID Q27404 standard; DNA; 34 BP.
 AC Q27404:
 DT 27-JAN-1993 (first entry)
 DE Human chromosome 21 centromere-specific L1.26 primer L126N.
 KW Polymerase chain reaction; PCR; repetitive element;
 KW Kpn family; Alu-like repeat family; ss.
 OS Synthetic.
 PN WO9213101-A.
 PD 06-AUG-1992.
 PF 24-JAN-1997; NL00018.
 PF 25-JAN-1991; NL-000132.
 PA (INGE-) INGENY BV.
 PI Uitterlinden AG, V1jg J;
 DR WPI: 92-284683/34.
 PT Detection of genetic variation by 2-D electrophoresis of
 PT fragments - and hybridisation with labelled probes, carried out
 PT on fragments consisting of inter-repeat sequences generated by
 PT PCR.
 PS Claim 13; Page 10; 31pp; English.
 CC Primer L126N is one of several primers which are preferred for use
 CC in amplifying inter-repeat regions of DNA. The amplified fragments
 CC are subjected to 2-D electrophoresis on the basis of length and
 CC differences in base sequence. The resulting separation pattern is
 CC transferred to a filter for screening with a probe. The method can
 CC be used to detect genetic variation.
 CC See Q27389-Q27404 and Q23141-Q23144.
 SQ Sequence 34 BP; 9 A; 11 C; 11 G; 3 T;

Query Match 100.0%; Score 9; DB 4; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 ttctcagaa 32
 |||||
 CP 33 ttctcagaa 25

```

RESULT 29
ID T69419 standard; DNA: 40 BP.
AC T69419;
DE 29-JUL-1997 (first entry)
Plasmid p1825f11 component oligonucleotide R18.
KW Oligonucleotide; plasmid: p1825f11; gene: promoter; R-Tem1;
beta-lactamase; construction: assembly; synthesis: end;
KW complementary; polymerase; reaction; EPCR; ss.
OS Synthetic.
PN WO9633207-A1.
PD 24-OCT-1996.
PF 18-APR-1996; U05480.
PR 18-APR-1995; US-425684.
PA (GLAXO) GLAXO GROUP LTD.
PI Liphutz RJ, Stemmer WPC;
WP1: 96-485725/46.
DR Polynucleotide amplification using bivalent PCR primer to generate
PT rolling circle intermediate or linear concatamers - requires reduced
PR amt. of primer compared to conventional PCR
PS Example 2; Fig 5A: 77pp; English.
CC The present sequence is a component oligonucleotide of the 2.71 kb
CC plasmid p1825f11 (Stemmer (1994) Nature 370:389), which encodes the
CC gene and promoter region for R-Tem1 beta-lactamase. A collection of
CC 134 component oligonucleotides, including the present sequence,
CC were synthesised and used to assemble the plasmid by end
CC complementary polymerase reaction (ECPRI), employing the overlapping
CC ends of the oligonucleotides. 10 C; 10 G; 9 T;
SO Sequence 40 BP; 11 A; 11 C; 10 G; 9 T;

Query Match 100.0%; Score 9; DB 30; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 14 ttctgagaa 22
OY 25 ttctgagaa 33

RESULT 30
ID Q37535 standard; DNA: 41 BP.
AC Q37535;
DE 17-JUN-1993 (first entry)
DE Shigella probe 1683.
KW Chromosome; Shigella; sonnei; flexneri; probe; Enteroinvasive; EIEC;
KW E. coli; transposable element; virulence plasmid; detection; dysentery;
KW NT11-2; ss.
OS Synthetic.
PN WO9303187-A.
PD 18-FEB-1993.
PF 28-JUL-1992; U06617.
PR 31-JUL-1991; US-738800.
PA (STRAD) AMOCO CORP.
PI McCarthy JW, Parodos K;
WP1: 93-076542/09.
DR Nucleic acid probes for detection of shigella and other pathogens
PT - used to diagnose dysentery in non-isotopic test format and have
PT utility in non-isotopic test formats requiring amplification for
PT high sensitivity.
PS Claim 35; Page 99; 129pp; English.
CC The sequences given in Q37534-45 are probes which were derived from
CC Shigella specific fragment NT11-2. These probes substantially retain
CC the exclusivity behaviour of NT11-2 towards non-enteroinvasive E. coli
CC (EIEC) Enterobacteriaceae. Probes such as this one can be used in the
CC detection of dysentery causing microbes. See also Q37506-35.
SO Sequence 41 BP; 11 A; 11 C; 8 G; 11 T;

Query Match 100.0%; Score 9; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 20 ttctgagaa 28
OY 25 ttctgagaa 33

```

```

RESULT 31
ID V54906 standard; DNA: 42 BP.
AC V54906;
DE 19-NOV-1998 (first entry)
DE Primer used to introduce E104K mutation into TEM-1 beta-lactamase gene.
KW TEM-1 beta-lactamase gene; recombination; gene shuffling; PCR primer; ss.
OS Synthetic.
PN US5811238-A.
PD 22-SEP-1998.
PF 30-NOV-1995; 564955.
PR 04-MAR-1996; US-537874.
PR 17-FEB-1994; US-198431.
PR 30-NOV-1995; US-564955.
PA (AFRY-) AFFIMAX TECHNOLOGIES NV.
PI Cramer A, Stemmer WPC;
WP1: 98-530860/45.
DR PCR primers V54900-14 were used to introduce mutations into the TEM-1
PT production of polynucleotides with desired properties - by iterative
PT selection and recombination
PS Example 6; Columns 49-50: 74pp; English.
CC PCR primers V54900-14 were used to introduce mutations into the TEM-1
CC beta-lactamase gene. The PCR products were used to construct mutant
CC combinations of the gene, using the method of the invention. The
CC specification describes a method for evolving a polynucleotide for
CC acquisition of a desired property. The method comprises providing a
CC population of variants of the polynucleotide, at least one of which is
CC in cell-free form, shuffling the variants of the polynucleotide to form
CC recombinant polynucleotides, selecting or screening for recombinant
CC polynucleotides that have evolved toward the desired property and
CC repeating the steps with the selected recombinant polynucleotides until
CC a recombinant polynucleotide has acquired the desired property. The
CC recombinant polynucleotide produced by method, or its expression product,
CC can be formulated as a pharmaceutical. 7 G; 12 T;
SO Sequence 42 BP; 14 A; 9 C; 9 G; 12 T;

Query Match 100.0%; Score 9; DB 49; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 ttctcagaa 11
CP 33 ttctcagaa 25

RESULT 32
ID T73035 standard; DNA: 42 BP.
AC T73035;
DE 12-FEB-1998 (first entry)
DE Mutagenic PCR primer 7 for the TEM-1 beta-lactamase gene.
KW TEM-1 beta-lactamase gene; gene reassembly reaction; DNA library;
KW gene shuffling; random fragmentation; mutagenesis; recombination;
KW GFP protein; arginate detoxification bacteria; cefotaxime;
KW calcium detoxification bacteria; drug resistance gene; PCR primer; ss.
OS Synthetic.
PN WO9720078-A1.
PD 05-JUN-1997.
PF 02-DEC-1996; U19256.
PR 25-MAR-1996; US-621859.
PR 30-NOV-1995; US-564955.
PA (AFRY-) AFFIMAX TECHNOLOGIES NV.
PI Cramer A, Stemmer WPC;
WP1: 97-310628/28.
DR Generating polynucleotide(s) with desired characteristics by
PT iterative selection and recombination - used for the directed
PT molecular evolution in vitro or in vivo of proteins, especially
PT green fluorescent protein
PS Example 6; Page 121; 209pp; English.
CC Mutagenic PCR primers T73029-43 were used to amplify the TEM-1
CC beta-lactamase gene, and introduce mutations into it. The present primer
CC introduces the amino acid change Glu104Lys. The TEM-1 beta-lactamase gene
CC confers resistance to bacteria against approximately 0.02 micrograms of
CC the drug cefotaxime. TEM-1 beta-lactamase is a very efficient enzyme.

```

CC limited in its reaction rate by diffusion. The mutated TEM-1
CC - beta-lactamase genes were used in a novel method to construct mutant
CC combinations of the gene, which produce enzymes with a greater activity.
CC The method involves recombining at least 2 forms of the polynucleotide to
CC produce a library of recombinant forms of the sequence. The library is
CC screened for a recombinant sequence, which is then recombined with
CC another form of the polynucleotide, the same or different from the first
CC and second forms, to produce a further library of recombinant
CC polynucleotides. The subsequent libraries are screened, and recombination
CC performed until the recombinant polynucleotide has acquired the desired
CC property. The methods are used to reassemble DNA after random
CC fragmentation, for mutagenesis of nucleic acid sequences by in vitro or
CC in vivo recombination. The repeated cycles of mutagenesis, shuffling and
CC selection allow for the directed molecular evolution in vitro or in vivo
CC of proteins. In particular the methods are used for the production of
CC mutant GFP protein which has been modified to show enhanced fluorescence
CC when used as a reporter of gene expression and regulation. Also
CC exemplified is the use of the methods to generate improved arsenate or
CC cadmium detoxification bacteria, drug resistance genes and variant
CC polymerases
SO Sequence 42 BP; 14 A; 9 C; 7 G; 12 T;
Query Match 100.0%; Score 9; DB 35; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3 tttcagaa 11
|||
Cp 33 tttcagaa 25
RESULT 33
ID N91955 standard; DNA; 50 BP.
AC N91955;
DT 13-APR-1990 (first entry)
DE Complementary strand of the Neisseria gonorrhoeae Tem-1 beta lactamase
DE gene combined with 11a2c amplifier probe
KW Neisseria gonorrhoeae; beta lactamase Tem-1; capture probe;
KW amplifier probe 11a2c; temkit21.18; TEM-1NH assay.
FH Neisseria gonorrhoeae
FS Key Location/Qualifiers
FT misc_feature 1..30
FT /*tag= a
FT /*capture probe"
FT misc_feature 31..50
FT /*tag= b
FT /*11a2c amplifier probe"
PN MO8903891-A.
PD 05-May-1989.
PE 14-OCT-1988: U03644.
PR 30-SEP-1988: US-252638, US-109282.
PI (CHIR-) Chiron Corp.
PI Udea MS, Warner B, Running JA, Kolberg JA, Clyne JM, Sanchez-Pescador R;
DR WPI: 89-150787/20.
PT Nucleic acid multimer for hybridisation assays
PT - having single-stranded oligo-nucleotide units
PT - capable of binding specifically to sequences of interest.
PS Fig. 10-2; 11pp; English.
CC Partial nucleotide sequences of the capture and amplifier probes used in
CC the TEM-1NH assay. The capture probe is complementary to the coding
CC strand of N. gonorrhoeae Tem-1 beta lactamase gene. It is called
CC temkit21.18.
SO Sequence 50 BP; 11 A; 11 C; 15 G; 13 T;
Query Match 100.0%; Score 9; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 5 tttcagaa 13
|||
Cp 25 tttcagaa 33

RESULT 34
ID V98966 standard; RNA; 50 BP.
AC V98966;
DT 17-MAR-1999 (first entry)
DE Human EGF-R hairpin ribozyme nucleotide, position 4701.
KW Human; epidermal growth factor receptor; EGFR; EGF-R; target sequence;
KW hammerhead ribozyme; hairpin ribozyme; inhibition; cell proliferation;
KW cancer; genetic drift; detection; mutation; ss.
OS Synthetic.
OS Homo sapiens.
PN WO9833893-A2.
PE 06-AUG-1998.
PF 14-JAN-1998: U00730.
PR 04-DEC-1997: US-0985162.
PR 31-JAN-1997: US-036476.
PA (RIBO-) RIBOZYME PHARM INC.
PI (UYAS-) UNIV ASTON.
PI Akhtar S, Fell P, McSwigen JA;
DR WPI: 98-437449/37.
PT Enzymatic nucleic acids - which cleave RNA derived from an epidermal
PT growth factor receptor, useful for inhibiting cell proliferation and
PT for treating cancers
PS Claim 7; Page 90; 109pp; English.
CC The present invention describes enzymatic nucleic acid molecules (NAMS)
CC which specifically cleave RNA derived from an epidermal growth factor
CC receptor (EGF-R) gene. V97221 to V98043 and V98979 to V99090 represent
CC specifically claimed target sequence from human EGF-R. V98044 to V98866
CC and V98867 to V9878 represent hammerhead ribozymes and hairpin ribozymes
CC respectively for human EGF-R. The NAMS are useful for cleaving EGF-R RNA
CC in the treatment of a condition associated with EGF-R expression levels
CC e.g. to inhibit cell proliferation in the prevention or treatment of
CC cancers. The NAMS can also be used as diagnostic tools to examine
CC genetic drift and mutations within diseased cells or to detect the
CC presence of EGF-R RNA in a cell.
SO Sequence 50 BP; 16 A; 10 C; 12 G; 12 U;
Query Match 100.0%; Score 9; DB 60; Length 50;
Best Local Similarity 66.7%; Pred. No. 1.37e+03;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 2 uucucagaa 10
|||
Cp 33 tttcagaa 25
RESULT 35
ID T99091 standard; DNA; 54 BP.
AC T99091;
DT 25-MAR-1998 (first entry)
DE N-terminal primer for ED7-HIV p17/p24 gene.
KW PCR primer; amplification; beta-galactosidase alpha region; ED7;
KW HIV protease recognition site; cross-linked enzyme donor polypeptide;
KW viral protease detection; ss.
OS Human immunodeficiency virus type 1.
PN WO9727294-A1.
PD 31-JUL-1997.
PE 23-JAN-1997.
PR 26-JAN-1996: US-592013.
PI (BOEF) BOEHRINGER MANNHEIM CORP.
PI Eisenbels SJ, Khanna P, Lingenfelter D, Manning WB,
PI Powell MJ, Tietze LF;
DR WPI: 97-393682/36.
PT Cross-linked enzyme donor polypeptide - which has reduced
PT complementation activity with enzyme acceptor polypeptide and may be
PT used in assays for e.g. viral protease
PS Example 5; Page 20; 36pp; English.
CC This sequence represents an amplification primer for a beta-galactosidase
CC alpha region (ED7) linked to a 8 residue (p17/p24) HIV protease
CC recognition site. The amplified sequence encodes a cross-linked enzyme
CC invention. The polypeptide of the invention is a cross-linked enzyme
CC donor polypeptide having an amino acid sequence substantially identical
CC to the N-terminal or C-terminal of native beta-galactosidase is. The
CC amino acid sequence has 2 cysteine residues intramolecularly linked via a

CC disulphide bond, or has 2 reactive amino acid residues intramolecularly
 CC linked via covalent attachment of the residues to a cross-linking agent,
 CC such that complementation of the enzyme donor polypeptide with an enzyme
 CC acceptor polypeptide is inhibited. The enzyme donor polypeptides may be
 CC used in assays, e.g. for viral proteases (see WO9727320). Such assays can
 CC be carried out with extremely low inherent background signal.

SO Sequence 54 BP; 14 A; 14 C; 10 G; 16 T;

Query Match 100.0%; Score 9; DB 36; Length 54;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 ttctcagaa 17
 |||||

Cp 33 ttctcagaa 25

RESULT 36
 ID V94164 standard; RNA; 54 BP.

AC V94164;
 DT 24-FEB-1999 (first entry)
 DE Human IL-2 receptor 9-chain hairpin ribozyme position 1419.
 KW Human; IL-2 receptor 9-chain; interleukin 2 receptor gamma chain;
 KW hammerhead ribozyme; hairpin ribozyme; substrate; expression; cancer;
 KW autoimmune disease; psoriasis; allergy; inflammatory disease;
 KW graft rejection; ss.

OS Synthetic.
 PN Homo sapiens.
 PN M09824913-A2.

PD 11-JUN-1998.
 PF 02-DEC-1997; U21748.
 PR 03-DEC-1996; US-758306.

PA (RIBO-) RIBOZYME PHARM INC.
 PI MCSwigen JA. Stinchcomb DT;

DR WPI; 98-33332/29.

PT Ribozymes targeted to interleukin 2 - useful for treating e.g.
 PT cancer, autoimmune disease and allergies

PS Claim 11; Page 39; 61pp; English.

CC The present sequence invention describes ribozymes targeted to modulate
 CC the synthesis and/or expression of interleukin (IL)-2R gamma encoded
 CC RNA. V93889 to V94574 represent specifically claimed ribozymes, and
 CC V94575 to V95260 represent specifically claimed substrate sequences
 CC from the present invention. The ribozymes can be used for the treatment
 CC of, e.g. graft rejection, autoimmune disease, cancer, psoriasis,
 CC allergy and other inflammatory conditions. The ribozymes are also used
 CC to induce tolerance in a recipient to alloantigen from a donor.

SO Sequence 54 BP; 17 A; 11 C; 11 G; 15 U;

Query Match 100.0%; Score 9; DB 54; Length 54;
 Best Local Similarity 66.7%; Pred. No. 1.37e+03;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 4 ttctcagaa 12
 ::|||

Cp 33 ttctcagaa 25

RESULT 37
 ID 062540 standard; DNA; 60 BP.

AC 062540;
 DT 06-DEC-1994 (first entry)
 DE Oligonucleotide probe for repeated sequence of D1721 locus.
 KW Probe; detection; repeat sequence; D1721; locus; control region;
 KW conserved sequence; human alpha satellite locus; mitochondria; ss.

OS Synthetic.

PN -EP-594959-A.

PD 04-MAY-1994.

PF 19-AUG-1993; 113261.

PR 28-AUG-1992; US-938084.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Reynolds RL. Walsh PS;

DR WPI; 94-145673/18.

PT Oligo:nucleotide probes for detecting human DNA - having

PT sequences complementary to a human alpha satellite locus or a
 PT conserved sequence in the mitochondrial genome

PS Claim 6; Page 17; 27pp; English.

CC This oligonucleotide probe is complementary to a region in the 2.7
 CC kilobase D1721 locus, a highly repetitive sequence. The use of
 CC probes complementary to a highly repeated sequence or to a conserved
 CC sequence provides high detection sensitivity and specificity. The
 CC probe may be optionally labelled with biotin and used in methods to
 CC quantitate the amount of DNA in a sample. See 062532-51.

SO Sequence 60 BP; 20 A; 10 C; 11 G; 19 T;

Query Match 100.0%; Score 9; DB 11; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 ttctcagaa 17
 |||||

Cp 33 ttctcagaa 25

RESULT 38
 ID 026744 standard; DNA; 63 BP.

AC 026744;
 DT 19-JAN-1993 (first entry)
 DE Epitope YOL1/34 coding sequence in PAP-1.
 KW Peptidoglycan-associated lipoprotein; polymerase chain reaction;
 KW tumour-associated antigen; antibody gene bank; screening;
 KW ribosome binding site; ss.

OS Synthetic.

PN Key

PN misc-feature

FT 4.57

FT /tag a

FT /note- "encodes TAG-linker

FT i.e. the YOL1/34 epitope"

PN DE4122598-C.

PD 30-JUL-1992.

PF 08-JUL-1991; 122598.

PR 08-JUL-1991; DE-122598.

PA (DEKR-) DEUT KREBSFORSCHUNGSZENT STIFTUNG.

PI Breitling F, Duebel S, Fuchs P, Little M;

DR WPI; 92-251202/31.

PT P-PSDB; R28523.

PT Plasmid PAP-1 and expression vectors for transforming bacterial

PT cell - contains recombinant DNA for encoding coupled Escherichia

PT coli surface antibody and PAL

PS Example 1; Fig 1b; 6pp; German.

CC The PAL gene from E.coli was amplified in plasmid pRC2 using the

CC polymerase chain reaction. The upstream primer introduces a BamHI

CC site at the 5' end of the gene resulting in the substitution of Cys

CC at position 1 of PAL by Gly. The amplified PAL sequence was ligated

CC to DNA fragments coding for the variable regions of heavy and light

CC chains of a humanised chicken lysozyme antibody joined via a DNA

CC fragment encoding an 18-amino acid sequence representing the epitope

CC YOL1/34 (the "TAG-linker"). Immediately upstream of the VH gene were

CC ligated a promoter/operator region, RBS and pelB leader sequence. The

CC resulting plasmid was designated PAP-1 and can be used for the rapid

CC screening of antibody gene banks, e.g. to isolate antibody against

CC tumour-associated antigen. See also 025907-8 and 026743-6.

SO Sequence 63 BP; 19 A; 15 C; 18 G; 11 T;

Query Match 100.0%; Score 9; DB 4; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 ttctcagaa 48
 |||||

Cp 33 ttctcagaa 25

RESULT 39

ID N60749 standard; DNA; 77 BP.

AC N60749;

DT 08-JUN-1991 (first entry)

```

DE Sequence encoding alpha-(Leu5)-endorphin.
KW Hormone: controlled pore glass: solid-phase synthesis; ss.
FH Key Location/Qualifiers
FT cds 11..64 /*tag= a
PN DE3544459-A.
PD 03-JUL-1986.
PR 16-DEC-1985: 544459.
PR 26-DEC-1984: JP-281645.
PA (NNSH ) NIPON SHINTAKU KK.
PI Yano J.
PI WPI: 86-177271/28.
DR P-PSDB: P60830.
PT Long-chain DNA e.g. endorphin gene prodn. - by solid-phase
PT coupling of oligonucleotide blocks on amnated controlled pore
PT glass CPG as carrier by tri:ester method
PS Example: page 8; 18pp; German.
CC The analgesic endocrine hormone endorphin was prepd. by synthesis of
CC the individual oligonucleotide blocks; linking of the individual
CC blocks by solid-phase synthesis on amnated CPG to give the gene;
CC synthesis of the double-stranded DNA and combination of it with a
CC vector plasmid; cloning of the resulting gene-contg. plasmid; and
CC culture of the transformed E. coli JM-103 strain produced. CPG is
CC "controlled pore glass".
SQ Sequence 77 BP; 17 A; 17 C; 20 G; 23 T;

Query Match 100.0%; Score 9; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 31 tctctagaa 39
OY 25 tctctagaa 33

RESULT 40
ID T79102 standard; RNA: 77 BP.
AC T79102:
DE RNA ligand to HIV-1 nucleocapsid.
KW Human immunodeficiency virus; HIV; nucleocapsid; SELEX; infection;
KM Inhibic; ss.
OS Synthetic.
PN US5635615-A.
PD 03-JUN-1997.
PR 11-JUN-1990: 536428.
PR 11-JUN-1990: US-536428.
PR 10-JUN-1991: US-714131.
PR 17-AUG-1992: US-931473.
PR 21-OCT-1992: US-964624.
PR 08-SEP-1993: US-117991.
PR 21-DEC-1994: US-361795.
PR 19-MAY-1995: US-447172.
PR 07-JUN-1995: US-477530.
PA (NEXS-) NEXSTAR PHARM INC.
PI Allen P, Gold L;
PI WPI: 97-309884/28.
PT Nucleic acids that are HIV-1 nucleocapsid ligands - for treatment of
PT diagnosis or HIV-1 infection
PS Claim 7; Column 21-22; 17pp; English.
CC T79101-T79122 are RNA ligands for the nucleocapsid of human
CC immunodeficiency virus type 1 (HIV-1). The ligands bind the
CC nucleocapsid and inhibit its function, they can be used in the
CC treatment and diagnosis of HIV-1 infection.
SQ Sequence 77 BP; 18 A; 22 C; 20 G; 17 U;

Query Match 100.0%; Score 9; DB 32; Length 77;
Best Local Similarity 66.7%; Pred. No. 1.37e+03;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 25 uucugagaa 33
OY 25 tctctagaa 33

```

```

RESULT 41
ID T21579 standard; cDNA to mRNA; 79 BP.
AC T21579:
DE 03-AUG-1996 (first entry)
DE Human gene signature HUNG502962.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994: J01916.
PR 12-NOV-1993: JP-355504.
PA (MARS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 917; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 79 BP; 18 A; 19 C; 12 G; 30 T;

Query Match 100.0%; Score 9; DB 19; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 70 tctctagaa 78
OY 25 tctctagaa 33

RESULT 42
ID N60751 standard; DNA: 80 BP.
AC N60751:
DE 08-JUN-1991 (first entry)
DE Sequence encoding gamma endorphin.
KW Hormone; controlled pore glass; solid-phase synthesis; ss.
FH Key Location/Qualifiers
FT cds 11..70 /*tag= a
PN DE3544459-A.
PD 03-JUL-1986.
PR 16-DEC-1985: 544459.
PR 26-DEC-1984: JP-281645.
PA (NNSH ) NIPON SHINTAKU KK.
PI Yano J.
PI WPI: 86-177271/28.
DR P-PSDB: P60832.
PT Long-chain DNA e.g. endorphin gene prodn. - by solid-phase
PT coupling of oligonucleotide blocks on amnated controlled pore
PT glass CPG as carrier by tri:ester method
PS Example: page 8; 18pp; German.
CC The analgesic endocrine hormone endorphin was prepd. by synthesis of
CC the individual oligonucleotide blocks; linking of the individual

```

CC blocks by solid-phase synthesis on aminated CPG to give the gene;
 CC synthesis of the double-stranded DNA and combination of it with a
 CC vector plasmid; cloning of the resulting gene-conv. plasmid; and
 CC culture of the transformed E. coli JM-103 strain produced. CPG is
 CC "controlled pore glass".
 SQ Sequence 80 BP; 15 A; 20 C; 22 G; 23 T;

Query Match 100.0%; Score 9; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Db 34 ttctcagaa 42
 |||||
 Oy 25 ttctcagaa 33

RESULT 43
 ID 051442 standard; DNA; 80 BP.
 AC 051442;
 DE 20-MAY-1994 (first entry)
 DE Human FACC Intron 7 3' region.
 KM Fanconi Anemia Group C; FACC; complementing CDNA; variant; diagnosis;
 KM open reading frame; Fanconi anemia; gene therapy; ss.
 OS Homo sapiens.
 PN W09322435-A.
 PD 11-NOV-1993.
 PF 27-APR-1993; CA0178.
 PR 29-APR-1992; US-876285.
 PR 21-JUL-1992; US-918313.
 PR 15-JAN-1993; US-003963.
 PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
 PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
 PI Buchwald M, Mathew CG, Strachee CA, Weirick R;
 DR WPI: 93-368794/46.
 PT Human CDNA which complements Fanconi Anaemia gp. C - used to
 PT develop prods. for use in diagnosis, study and therapy of Fanconi
 PT Anemia
 PS Claim 1; Page 11; 137pp; English.
 CC The sequences given in 051429-54 represent the 5' and 3' splice
 CC regions of the introns from the Fanconi Anemia Group C complementing
 CC (FACC) DNA. Three CDNA molecules which are cellular variants of a
 CC single CDNA, are transcribed from the FACC gene. The three cDNAs
 CC each contain an identical open reading frame encoding the FACC
 CC protein. FACC protein may be used for the diagnosis and study of
 CC Fanconi anemia. The FACC gene and cDNAs may be used in gene therapy.
 SQ Sequence 80 BP; 19 A; 15 C; 14 G; 31 T;

Query Match 100.0%; Score 9; DB 9; Length 80;
 Best Local Similarity 88.9%; Pred. No. 1.37e+03; Indels 0; Gaps 0;
 Matches 8; Conservative 1; Mismatches 0;

Db 27 ttctcagaa 35
 |||||
 Oy 25 ttctcagaa 33

RESULT 44
 ID 062542 standard; DNA; 100 BP.
 AC 062542;
 DE 06-DEC-1994 (first entry)
 DE Oligonucleotide probe for repeated sequence of D17Z1 locus.
 KM Probe; detection; repeat sequence; D17Z1; locus; control region;
 KM conserved sequence; human alpha satellite locus; mitochondria; ss.
 OS Synthetic.
 PN EP-594959-A.
 PD 04-MAY-1994.
 PF 19-AUG-1993; 113261.
 PR 28-AUG-1992; US-938084.
 PA (HOFF-) HOFFMANN LA ROCHE & CO AG F.
 PI Reynolds RL, Walsh PS;
 DR WPI: 94-145673/18.
 PT Oligonucleotide probes for detecting human DNA - having
 PT sequences complementary to a human alpha satellite locus or a

PT conserved sequence in the mitochondrial genome
 PT Claim 6; Page 18; 27pp; English.
 CC This oligonucleotide probe is complementary to a region in the 2.7
 CC kilobase D17Z1 locus, a highly repetitive sequence. The use of
 CC probe complementary to a highly repeated sequence or to a conserved
 CC sequence provides high detection sensitivity and specificity. The
 CC probe may be optionally labelled with biotin and used in methods to
 CC quantitate the amount of DNA in a sample. See 062532-51.
 SQ Sequence 100 BP; 30 A; 18 C; 20 G; 32 T;

Query Match 100.0%; Score 9; DB 11; Length 100;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Db 9 ttctcagaa 17
 |||||
 Cp 33 ttctcagaa 25

RESULT 45
 ID 062537 standard; DNA; 130 BP.
 AC 062537;
 DE 06-DEC-1994 (first entry)
 DE Oligonucleotide probe for repeated sequence of D17Z1 locus.
 KM Probe; detection; repeat sequence; D17Z1; locus; control region;
 KM conserved sequence; human alpha satellite locus; mitochondria; ss.
 OS Synthetic.
 PN EP-594959-A.
 PD 04-MAY-1994.
 PF 19-AUG-1993; 113261.
 PR 28-AUG-1992; US-938084.
 PA (HOFF-) HOFFMANN LA ROCHE & CO AG F.
 PI Reynolds RL, Walsh PS;
 DR WPI: 94-145673/18.
 PT Oligonucleotide probes for detecting human DNA - having
 PT sequences complementary to a human alpha satellite locus or a
 PT conserved sequence in the mitochondrial genome
 PS Claim 6; Page 15; 27pp; English.
 CC This oligonucleotide probe is complementary to a region in the 2.7
 CC kilobase D17Z1 locus, a highly repetitive sequence. The use of
 CC probes complementary to a highly repeated sequence or to a conserved
 CC sequence provides high detection sensitivity and specificity. The
 CC probe may be optionally labelled with biotin and used in methods to
 CC quantitate the amount of DNA in a sample. See 062532-51.
 SQ Sequence 130 BP; 32 A; 26 C; 28 G; 44 T;

Query Match 100.0%; Score 9; DB 11; Length 130;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Db 9 ttctcagaa 17
 |||||
 Cp 33 ttctcagaa 25

RESULT 46
 ID T03931 standard; CDNA; 149 BP.
 AC T03931;
 DE 20-DEC-1995 (first entry)
 DE Human Factor-V intron 10.
 KM Factor-V; thrombosis; thrombophilia; diagnosis; anticoagulant;
 KM activated protein-C; APC; homozygosity; heterozygosity; ss.
 OS Homo sapiens.
 PN W09521938-A1.
 PD 17-AUG-1995.
 PF 14-FEB-1995; E00553.
 PR 14-FEB-1994; EP-200377.
 PA (OYLE-) RICKSUNIV LEIDEN.
 PI Bertina RM, Reitsma PH;
 DR WPI: 95-293134/38.
 PT Screening for genetic defect associated with thrombosis and/or poor
 PT anticoagulant response to activated protein C - useful to determine
 PT homozygosity or heterozygosity for a mutation in Factor V, Va, VIII

PT or VIIA.
 PS - Disclosure: Page 70-71: 98bp; English.
 CC Human Factor-V Intron 10 DNA was generated by hemi-nested PCR.
 CC From this sequence, primers were designed for the amplification
 CC of 2 overlapping genomic fragments that can be used for
 CC genotyping, to detect the 1691A allele associated with APC
 CC resistance.
 SO Sequence 149 BP; 41 A; 26 C; 29 G; 53 T;
 Query Match 100.0%; Score 9; DB 15; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 32 tttcagaa 40
 25 tttcagaa 33
 OY

RESULT 47
 ID 061212 standard; DNA; 195 BP.
 AC 061212:
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST01247.
 KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcription; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN M09316178-A.
 PD 19-AUG-1993:
 PF 12-FEB-1993: U01294.
 PR 12-FEB-1992: US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ;
 DR WPI: 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 PS Example 4; Page 455; 500bp; English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prepn. of antisense sequences, probes and constructs.
 CC EST01247 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CRM. See also Q59041-Q61440.
 SO Sequence 195 BP; 51 A; 53 C; 50 G; 37 T;
 Query Match 100.0%; Score 9; DB 8; Length 195;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 12 tttcagaa 20
 33 tttcagaa 25
 Cp

RESULT 48
 ID 030026 standard; DNA; 197 BP.
 AC 030026:
 DT 18-MAR-1993 (first entry)
 DE PCR product prepared using primer WBS2.
 KW PCR; chromosome specific repeated DNA; chromosome staining;
 KW cytogenetics;interphase nuclei; metaphase spreads; germ line cells;
 KW somatic cells; ss.
 OS Synthetic.
 PN EP-511750-A.
 PD 04-NOV-1992:
 PF 09-APR-1992: 303159.
 PR 09-APR-1991: US-683441.
 PR 26-MAR-1992: US-858124.
 PA (RESC) UNIV CALIFORNIA.
 PI Gray JW, Welter HUG;

DR WPI: 92-367577/45.
 PT Deoxyribonucleic acid amplification using degenerate oligo
 PT nucleotide primers - useful for chromosome specific repeated DNA
 PT for staining agent in cyto genetic analysis by polymerase chain
 PT reaction
 PS Claim 16; Page 18; 25pp; English.
 CC This sequence represents that of the insert of plasmid pBS609-51
 CC created by the ligation of the PCR product prepared using primer
 CC WBS2 (Q30017) into pBS. PCR primer Q30017 is used to amplify a
 CC region of the 171 bp alpha satellite (alphoid) repeat sequence,
 CC conserved in all human chromosomes. This binds to the alpha
 CC satellite repeat consensus sequence. The minimal product size
 CC expected is 175bp, and is used as a probe for the repeat sequence
 CC of human centromeres.
 SO Sequence 197 BP; 61 A; 37 C; 43 G; 56 T;
 Query Match 100.0%; Score 9; DB 5; Length 197;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 158 tttcagaa 166
 33 tttcagaa 25
 Cp

RESULT 49
 ID 030027 standard; DNA; 197 BP.
 AC 030027:
 DT 18-MAR-1993 (first entry)
 DE PCR product prepared using primer WBS4.
 KW PCR; chromosome specific repeated DNA; chromosome staining;
 KW cytogenetics;interphase nuclei; metaphase spreads; germ line cells;
 KW somatic cells; ss.
 OS Synthetic.
 PN EP-511750-A.
 PD 04-NOV-1992:
 PF 09-APR-1992: 303159.
 PR 09-APR-1991: US-683441.
 PR 26-MAR-1992: US-858124.
 PA (RESC) UNIV CALIFORNIA.
 PI Gray JW, Welter HUG;
 DR WPI: 92-367577/45.
 PT Deoxyribonucleic acid amplification using degenerate oligo
 PT nucleotide primers - useful for chromosome specific repeated DNA
 PT for staining agent in cyto genetic analysis by polymerase chain
 PT reaction
 PS Claim 16; Page 18; 25pp; English.
 CC This sequence represents that of the insert of plasmid pBS609-52
 CC created by the ligation of the PCR product prepared using primer
 CC WBS4 (Q30018) into pBS. PCR primer Q30018 is used to amplify a
 CC region of the 171 bp alpha satellite (alphoid) repeat sequence,
 CC conserved in all human chromosomes. This binds to the alpha
 CC satellite repeat consensus sequence. The minimal product size
 CC expected is 175bp, and is used as a probe for the repeat sequence
 CC of human centromeres.
 SO Sequence 197 BP; 60 A; 35 C; 42 G; 60 T;
 Query Match 100.0%; Score 9; DB 5; Length 197;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 158 tttcagaa 166
 33 tttcagaa 25
 Cp

RESULT 50
 ID T79775 standard; DNA; 198 BP.
 AC T79775:
 DT 10-NOV-1997 (first entry)
 DE Potato starch branching enzyme; SBE; potato; genetic engineering;
 KW starch branching enzyme; SBE; potato; genetic engineering;
 KW transgenic plant; starch production; modification; ss.

OS Solanum tuberosum.
 PN M0970413-A2.
 PD 06-FEB-1997.
 PF 12-JUL-1996; E03053.
 PR 14-JUL-1995; GB-014437.
 PA (DANI-) DANISCO AS.
 PI Poulsen P.
 DR WPI: 97-132651/12.
 PT Affecting enzymatic activity in plant by expressing sense intron
 sequence - esp. for inhibiting starch branching enzyme to create
 modified starch without post-harvest derivatization
 PS Claim 8; Page 31; 70pp; English.
 CC T79775 represents intron 6 of the potato starch-branching enzyme
 (SBE) gene. Sequences antisense to introns of the SBE gene are used
 to produce transgenic plants producing modified or new forms of starch
 for industrial use. The sequences allow starches to be made without
 post-harvest modification (which uses hazardous chemicals and requires
 much energy). Using the antisense intron sequences is more reliable,
 efficient and selective than known methods for regulating enzymatic
 activity because there is almost no homology between introns in
 different SBE genes. The SBE gene promoter (see T79783) is also
 useful in genetic engineering.
 CC Sequence 198 BP; 64 A; 32 C; 29 G; 73 T;
 SO Query Match 100.0%; Score 9; DB 33; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 32 ttctcagaa 40
 Cp 33 ttctcagaa 25

RESULT 51
 ID T22515 standard; cDNA to mRNA; 209 BP.
 AC T22515;
 DT 02-SEP-1996 (first entry)
 DE Human gene signature HUMG504127.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W0514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATS/) MATSUBARA K.
 PI (OKUBO/) OKUBO K.
 PI Matsubara K; Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1; Page 1447; 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 783 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridize to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SO Sequence 209 BP; 51 A; 31 C; 40 G; 86 T;
 Query Match 100.0%; Score 9; DB 20; Length 209;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 96 ttctcagaa 104
 Cp 25 ttctcagaa 33

RESULT 52
 ID Q77502 standard; DNA; 218 BP.
 AC Q77502;
 DT 23-SEP-1994 (first entry)
 DE Human genome fragment. (Preferred)
 KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
 KW detection; homology; human; adrenal tissue; ds.
 OS Homo sapiens.
 PN W09401346-A.
 PD 20-JAN-1994.
 PF 13-JUL-1993; G01467.
 PR 13-JUL-1992; GB-014857.
 PA (MEDI-) MEDICAL RES COUNCIL.
 PI Gross J; Hadfield KM; Howells D; Kelly M; Shaw D;
 PI Sibson DR; Starkey M;
 DR WPI: 94-035056/04.
 PT New nucleic acid fragment encoding gene products - can be used
 PT for genetic analysis and mapping
 PS Claim 1; Page 560; 616pp; English.
 CC Human nucleic acid fragments, isolated from brain, adrenal tissue,
 CC the placenta or bone marrow comprise any of: (A) a sequence
 CC selected from (Q76401-Q77613), (B) an allelic variation of a
 CC sequence as described in (A), or (C) a sequence complementary
 CC to (A) or (B).
 CC Preferred sequences exhibit no more than 90% homology to a human
 CC sequence known per se.
 SO Sequence 218 BP; 66 A; 52 C; 43 G; 48 T;
 Query Match 100.0%; Score 9; DB 10; Length 218;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 105 ttctcagaa 113
 Cp 33 ttctcagaa 25

RESULT 53
 ID Q73722 standard; DNA; 243 BP.
 AC Q73722;
 DT 18-JUL-1995 (first entry)
 DE Retrotransposon probe Tos12, cloned from Oryza sativa.
 KW Retrotransposon; probe; rice; plant; identification; genetic uniformity;
 KW pure line; inbred parent line; near-isogenic line; ss.
 OS Oryza sativa.
 PN J06277063-A.
 PD 04-OCT-1994.
 PF 08-OCT-1992; 309167.
 PR 08-OCT-1992; JP-309167.
 PA (NORI-) NORIN SUISANSHO NOGYO SEIBUTSU SHIGEN.
 PA (TAKE-) TAKEDA CHEM IND LTD.
 DR WPI: 94-353746/44.
 PT Identification of plants at the DNA level - using retrotransposon
 PT sequences as probes
 PS Claim 6; Page 10; 41pp; Japanese.
 CC The nucleotide sequence of the retrotransposon probe Tos12 cloned from
 CC the rice plant Oryza sativa. The probes may be used for the analysis of
 CC plant genes to secure the efficient identification of plants and genetic
 CC uniformity for the production of pure lines, inbred parent lines and
 CC near-isogenic lines.
 SO Sequence 243 BP; 67 A; 33 C; 57 G; 86 T;
 Query Match 100.0%; Score 9; DB 13; Length 243;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 162 ttctcagaa 170
 Cp 33 ttctcagaa 25

RESULT 54
 ID V88500 standard; cDNA; 251 BP.
 AC V88500;
 DT 12-FEB-1999 (first entry)
 DE EST clone HR693.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN WO9845437-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U06956.
 PR 10-APR-1997: US-837312.
 PA (GENE) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racine LA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1: Page 417: 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SO Sequence 251 BP; 60 A; 47 C; 39 G; 105 T;

Query Match 100.0%; Score 9; DB 53; Length 251;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 ttctcagaa 36
 Cp 33 ttctcagaa 25

RESULT 55
 ID T93999 standard; DNA; 258 BP.
 AC T93999;
 DT 24-FEB-1998 (first entry)
 DE Model pre-therapeutic RNA molecule.
 KW Diphtheria toxin; subunit A; human chorionic gonadotropin; HCG;
 KW prophetic hyperthermy; pre-therapeutic molecule; Gaucher's disease;
 KW functional therapeutic RNA; splicing; human immunodeficiency virus;
 KW autoimmune disease; tissue rejection; therapeutic protein; ss.
 OS Synthetic.
 FH Key
 FT misc_binding
 FT 1..22
 FT /tag- a
 FT /note- "Complementary to the target sequence of the
 adenovirus 2 major late promoter"
 FT 23..53
 FT /tag- b
 FT /note- "spacer region"
 FT 60
 FT misc_feature
 FT /tag- c

FT /note- "branch point adenosine"
 FT 61..83
 FT /tag- d
 FT /note- "contains polypyrimidine tract"
 FT 84..85
 FT /tag- e
 FT /note- "3' splice site"
 FT 86..127
 FT /tag- f
 FT /note- "HSV protein marker"
 FT 128..145
 FT /tag- g
 FT /note- "histidine protein marker"
 FT 151..153
 FT /tag- h
 FT /note- "STOP codon"

FT W09722250-A1.
 FT 26-JUN-1997.
 FT 13-DEC-1996: U19419.
 FT 15-DEC-1995: US-008717.
 FT (INTR-) INTRON LLC.
 FT Mitchell LG;
 FT WPI: 97-341349/31.
 PT New pre-therapeutic RNA converted to active form by trans-splicing -
 PT only in cells containing specific pre-mRNA, used to generate, e.g.
 PT toxins for targeted killing of cells such as tumour or
 PT virus-infected cells
 PS Example 2: Page 18: 47pp: English.
 CC The present sequence represents a model pre-therapeutic RNA molecule
 CC comprising a binding region for target pre-mRNA, at least one of 3' and
 CC 5' splice sites, sequences functional as RNA or sequences encoding a
 CC diphtheria toxin, see T93980), splicing modulation sequences and sequence
 CC elements required for splicing and production of a therapeutic RNA.
 CC The inclusion of the histidine protein marker allows for detection using
 CC metal chelation chromatography and the HSV protein marker allows for
 CC detection using monoclonal antibodies. Pre-therapeutic molecules are
 CC used to produce functional therapeutic RNA or nucleic acid analogues in
 CC specific target cells, which may be present in animals, plants, and
 CC especially humans. The pre-therapeutic molecule does not have to be
 CC targeted to a particular type of cell, as the therapeutic molecule
 CC is only produced in those cells that contain target pre-RNA. Typical
 CC products of the RNA produced by splicing are toxins that kill cells
 CC infected with virus (e.g. human immunodeficiency virus) or other
 CC pathogens, benign and malignant neoplastic cells (e.g. prostatic
 CC hyperthermy) or components of the immune system involved in autoimmune
 CC disease and tissue rejection. Otherwise the product generates a
 CC therapeutic protein, marker or new gene product not normally expressed
 CC in these cells, e.g. for treatment of inherited diseases (e.g. Gaucher's
 CC disease) or to correct for a missing/defective gene.
 SO Sequence 258 BP; 68 A; 72 C; 64 G; 54 T;

Query Match 100.0%; Score 9; DB 35; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 38 ttctgagaa 46
 Oy 25 ttctgagaa 33

RESULT 56
 ID T01037 standard; cDNA; 261 BP.
 AC T01037;
 DT 07-MAY-1996 (first entry)
 DE Mouse B7-2 exon 5.
 KW T-cell costimulatory molecule; B7-2; T-lymphocyte; CD28; CTLA4;
 KW counter receptor; transgenic animal; immunoglobulin; ds.
 OS Mus musculus.
 FH Key
 FT cds
 FT 1..138
 FT /tag- a
 FT W09523859-A2.

PD 08-SEP-1995.
 PF 02-MAR-1995: U02576.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PI (DAND) DNA FARMER CANCER INST.
 DR Bortelle F, Freeman GJ, Nadler LM, Sharpe AH:
 WPI: 95-320574/41.
 DR P-PSB: R82890.
 PT Novel T cell co-stimulatory molecules - corresponding to naturally
 occurring alternatively spliced forms of T cells co-stimulatory
 PT molecules or variants
 PS Claim 9: Page 80; 11pp: English.
 CC Exon 5 (701037) of the mouse T-cell costimulatory molecule B7-2
 CC gene encodes the cytoplasmic domain (R82890) of B7-2. It can
 CC be used to construct nucleic acids coding for novel cytoplasmic
 CC forms of T-cell costimulatory molecules that also incorporate a
 CC second cytoplasmic domain. Such T-cell costimulatory molecules bind
 CC to CD28 or CTLA4 and trigger a costimulatory signal in T-cells.
 SQ Sequence 261 BP; 97 A; 52 C; 59 G; 53 T;

Query Match 100.0%; Score 9; DB 17; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 201 ttctcagaa 209
 |||||||
 Cp 33 ttctcagaa 25

RESULT 57
 ID Q96134 standard; DNA: 282 BP.
 AC Q96134;
 DT 12-APR-1996 (first entry)
 DE T-cell receptor beta-chain variable region V-beta-14-encoding DNA.
 KW Diabetes; adoptive immunotherapy; gene therapy;
 KM T-cell receptor beta-chain; ss.
 OS Homo sapiens.
 PN MO9521623-A1.
 PD 17-AUG-1995.
 PF 10-FEB-1995: U01572.
 PR 14-FEB-1994: US-195963.
 PA (UYVE-) UNIV VERMONT.
 PI Albertini RJ, Falca MT;
 DR WPI: 95-292941/38.
 PT Preventing or reducing severity of diabetes - by inhibiting the
 PT activity of specific T-cells, partic. by interfering with
 PT diabetes-associated T cell receptors
 PS Disclosure; Fig 7; 42pp: English.
 CC DNA sequences (Q96132-34) coding for non-conserved regions of T-cell
 CC receptor (TCR) beta-chain variable regions V-beta-6.1, V-beta-6.6/6.7
 CC and V-beta-14 (R78686-88) may be operatively linked to expression
 CC control sequences and incorporated into vectors. These vectors
 CC are directly injected into e.g. the skeletal muscle of a diabetic
 CC patient and the encoded proteins cause an effect on the immune
 CC system that interferes with diabetes-associated TCRs. Diabetes
 CC is an autoimmune disease characterized by predominant usage of TCRs
 CC bearing V-beta-6 and V-beta-14.
 SQ Sequence 282 BP; 84 A; 67 C; 69 G; 62 T;

Query Match 100.0%; Score 9; DB 17; Length 282;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69 ttctcagaa 77
 |||||||
 Cp 33 ttctcagaa 25

RESULT 58
 ID V69007 standard; DNA: 293 BP.
 AC V69007;
 DT 22-JAN-1999 (first entry)
 DE DNA molecule encoding a polypeptide for detecting breast cancer #9.

KW Human: breast cancer; breast tumour tissue; diagnosis; treatment;
 KM vaccine; epitope; endogenous; retroviral element; ss.
 OS Homo sapiens.
 PN MO9845328-A2.
 PD 15-OCT-1998.
 PF 09-APR-1998: U06939.
 PR 11-DEC-1997: US-991789.
 PA 09-APR-1997: US-838762.
 PI (COR-) CORIXA CORP.
 PI Fridakis TN, Reed SG, Smith JM:
 DR WPI: 98-557473/47.
 PT New DNA sequences isolated from endogenous human retroviral element
 PT - and related vectors, transformed cells, proteins and antibodies,
 PT useful for diagnosis, treatment and prevention of breast cancer
 PS Claim 11: Page 65; 173pp: English.
 CC V68999 to V69038 represent nucleotide sequences which encode polypeptides
 CC used in detecting human breast cancer. Detection or measurement of
 CC human breast tumour specific polypeptides and nucleotide sequences,
 CC or the corresponding RNA in a sample, is used for diagnosis and
 CC monitoring of breast cancer. Human breast tumour specific polypeptides
 CC and nucleotide sequences, and the vectors containing the DNAs, are also
 CC useful in vaccines for inhibiting development (for prevention or
 CC therapy) of breast cancer. The polypeptides may also be used to
 CC raise monoclonal antibodies, used as immunoassay reagents.
 SQ Sequence 293 BP; 66 A; 43 C; 48 G; 73 T;

Query Match 100.0%; Score 9; DB 52; Length 293;
 Best Local Similarity 88.9%; Pred. No. 1.37e+03;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 176 ttctcagaa 184
 |||||||
 Cp 33 ttctcagaa 25

RESULT 59
 ID V87805 standard; CDNA: 296 BP.
 AC V87805;
 DT 12-FEB-1999 (first entry)
 DE EST clone EH61.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KM tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN MO9845437-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U06956.
 PR 10-APR-1997: US-837312.
 PA (GEM) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavalie ER, McCoy JM, Merberg D,
 PI Racine LA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1: Page 186; 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SQ Sequence 296 BP; 94 A; 57 C; 64 G; 81 T;

Query Match 100.0%; Score 9; DB 54; Length 296;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 115 ttctagaa 123
 |||||
 25 ttctagaa 33

Query Match 100.0%; Score 9; DB 3; Length 297;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 91 unucagaa 99
 |||||
 33 ttctagaa 25

RESULT 60
 ID N50452 standard; RNA: 297 BP.
 AC N50452:
 DT 12-MAR-1992 (first entry)
 DE HOP stunt viroid RNA.
 KW HSV; ss.
 OS HOP stunt viroid.
 PN J60041486-A.
 PD 05-MAR-1985.
 PF 17-AUG-1983: 149815.
 PR 17-AUG-1983: JP-149815.
 PA (KIRI) KIRIN BREWERY KK.
 DR WPI: 85-090909/15.
 PT New plasmid(s) with consecutive fractions contg. cDNA - corresp.
 PI to plant virus RNA, for introducing exogenous genes into plant
 PT cells etc.
 PS Disclosure: Fig 1: 14pp: Japanese.
 CC The RNA folds back on itself to form a stem loop. cDNA
 CC corresponding to the RNA can be incorporated into plasmids and
 CC these used for the introduction of exogenous genes into plants or
 CC plant cells.
 CC See also N50453 and N50454.
 SQ Sequence 297 BP; 61 A; 88 C; 79 G; 69 U;

Query Match 100.0%; Score 9; DB 3; Length 297;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 91 unucagaa 99
 |||||
 33 ttctagaa 25

RESULT 61
 ID V16717 standard; DNA: 297 BP.
 AC V16717:
 DT 15-JUN-1998 (first entry)
 DE Nucleotide sequence encoding the HPV-16 E7 oncoprotein.
 KW E7 oncoprotein; proliferative state; HPV; kinase activity;
 KW cyclin/cyclin-dependent kinase; p131p1; interaction; inactivation;
 KW cyclin/cyclin-dependent kinase inhibitor; ss.
 OS Human papillomavirus.
 PA Key
 FH Location/Qualifiers
 FT CDS
 FT 1..297
 FT /*tag= a

US5736318-A.
 07-APR-1998.
 17-MAR-1995: 406248.
 17-MAR-1995: US-406248.
 (HARD) HARVARD COLLEGE.
 PA (HARD) UNIV HARVARD.
 PI Jones DL, Munger K;
 DR WPI: 98-239202/21.
 DR P-PEDB: W46886.
 PT Evaluation of proliferative state of cells transformed with human
 PT papilloma virus - by determining cyclin-dependent kinase activity
 PT induced by E7 onco-protein
 PS Disclosure: Columns 19-20: 14pp: English.
 CC The present sequence encodes Human papillomavirus (HPV), strain 16,
 CC E7 oncoprotein. The proliferative state of a cell transformed with
 CC HPV can be evaluated in the following manner. Cyclin/cyclin-dependent
 CC kinase complexes containing protein p131p1 (W46886-88) are isolated
 CC from the transformed cell, and the HPV E7 oncoprotein added to the
 CC isolated protein. Cyclin/cyclin-dependent kinase complexes are isolated

CC from an untransformed cell that is substantially homogenic with the
 CC transformed cell, and the HPV E7 oncoprotein added. The kinase activities
 CC of the 2 samples are measured, where a proliferating transformed cell
 CC has a greater kinase activity than the untransformed cell. The method is
 CC used for determining the extent of interaction and/or inactivation
 CC between a cyclin/cyclin-dependent kinase inhibitor and the HPV E7
 CC oncoprotein and thus evaluating the proliferative state of a transformed
 CC cell.

Sequence 297 BP; 93 A; 61 C; 67 G; 76 T;

Query Match 100.0%; Score 9; DB 40; Length 297;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 282 ttctagaa 290
 |||||
 33 ttctagaa 25

Query Match 100.0%; Score 9; DB 10; Length 306;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 157 ttctagaa 165
 |||||
 33 ttctagaa 25

RESULT 62
 ID O57426 standard; cDNA to mRNA: 306 BP.
 AC O57426:
 DT 19-OCT-1994 (first entry)
 DE Zinc resistant like protein.
 KW Human cDNA; library; enzyme; protein; ss.
 OS Homo sapiens.
 PN W09403599-A.
 PD 17-FEB-1994.
 PF 04-AUG-1993: J01095.
 PR 04-AUG-1992: JP-208077.
 PR 13-NOV-1992: JP-327619.
 PR 26-FEB-1993: JP-061431.
 PA (SAGA) SAGAMI CHEM RES CENTRE.
 PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;
 DR WPI: 94-065688/08.
 DR P-PEDB: R46087.
 PT cDNA of human origin and proteins coded by it - which may be
 PT expressed by in vivo or in vitro translation using sense RNA or
 PT antisense DNA corresponding to the cDNA.
 PS Claim 1: Page 41-42: 167pp: Japanese.
 CC mRNA expressed in human fibrosarcoma cell line HT-1080 was
 CC isolated and used to construct a cDNA library using vector
 CC pK1. Clone HP00062 encoding Zn resistant-like protein
 CC was isolated.

Sequence 306 BP; 66 A; 57 C; 72 G; 111 T;

Query Match 100.0%; Score 9; DB 10; Length 306;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 157 ttctagaa 165
 |||||
 33 ttctagaa 25

RESULT 63
 ID O76420 standard; DNA: 314 BP.
 AC O76420:
 DT 23-SEP-1994 (first entry)
 DE Human genome fragment. (Preferred)
 KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
 KW detection; homology; human; adrenal tissue; ds.
 OS Homo sapiens.
 PN W09401548-A.
 PD 20-JAN-1994.
 PF 13-JUL-1993: G01467.
 PR 13-JUL-1992: GB-014857.
 PA (MEDIC) MEDICAL RES COUNCIL.
 PI Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;
 DR WPI: 94-035056/04.
 DR New nucleic acid fragment encoding gene products - can be used
 PT for genetic analysis and mapping

PS Claim 1: Page 31: 616pp: English.
 CC Human nucleic acid fragments, isolated from brain adrenal tissue,
 CC the placenta or bone marrow comprise any of: (A) a sequence
 CC selected from (076401-077613), (B) an allelic variation of a
 CC sequence as described in (A), or (C) a sequence complementary
 CC to (A) or (B).
 CC Preferred sequences exhibit no more than 90% homology to a human
 CC sequence known per se.
 SQ Sequence 314 BP: 105 A: 36 C: 41 G: 132 T:
 Query Match 100.0%; Score 9: DB 10: Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 140 ttctgagaa 148
 |||||||||
 QY 25 ttctgagaa 33

RESULT 64
 ID Q60687 standard: cDNA: 325 BP.
 AC Q60687;
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST00813.
 KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcription: mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN M09316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993: U01294.
 PR 12-FEB-1992: US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD. Moreno RF. Venter CJ.
 DR WPI: 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 PS Example 4: Page 374: 500pp: English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC EST00813 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CRM. See also Q59041-Q61440.
 SQ Sequence 325 BP: 100 A: 35 C: 60 G: 127 T:
 Query Match 100.0%; Score 9: DB 8: Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 102 ttctgagaa 110
 |||||||||
 QY 25 ttctgagaa 33

RESULT 65
 ID V87753 standard: cDNA: 326 BP.
 AC V87753;
 DT 12-FEB-1999 (first entry)
 DE EST clone EQ187.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST: human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN M09845437-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U06956.
 PR 10-APR-1997: US-837312.
 PA (GENM) GENETICS INST. INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D,
 PI Racie LA, Spaulding V, Treacy M.
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1: Page 167-168: 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SQ Sequence 326 BP: 77 A: 97 C: 75 G: 77 T:
 Query Match 100.0%; Score 9: DB 54: Length 326;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 50 ttctcagaa 58
 |||||||||
 CP 33 ttctcagaa 25

RESULT 66
 ID T21863 standard: cDNA to mRNA: 329 BP.
 AC T21863;
 DT 14-AUG-1996 (first entry)
 DE Human gene signature HUMGS03404.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN M09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994: J01916.
 PR 12-NOV-1993: JP-355504.
 PA (MATS/) MATSUBARA K.
 PA (OKUB/) OKUBO K.
 PI Matsubara K. Okubo K.
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 985-986: 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in 119001-126837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 329 BP: 124 A: 43 C: 59 G: 101 T:
 Query Match 100.0%; Score 9: DB 20: Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 10 ttctgagaa 18
 |||||
 Qy 25 ttctgagaa 33

RESULT 67
 ID Q59108 standard; cDNA; 329 BP.
 AC Q59108:
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST00070.
 KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcription; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN W0316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993: U01294.
 PR 12-FEB-1992: US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ;
 DR WPI: 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 PI markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 PS Example 1, Page 122: 500pp: English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC EST00070 has a "poor" coding probability as evaluated using the
 CC coding-region prediction program CRM.
 CC See also Q59041-061440.
 SQ Sequence 329 BP; 87 A; 74 C; 96 G; 69 T;

Query Match 100.0%; Score 9; DB 8; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 73 ttctgagaa 81
 |||||
 Qy 25 ttctgagaa 33

RESULT 68
 ID T26204 standard; cDNA to mRNA; 337 BP.
 AC T26204:
 DT 13-NOV-1996 (first entry)
 DE Human gene signature H0MGS08443.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994: J01916.
 PR 12-NOV-1993: JP-355504.
 PA (MATS/) MATSUBARA K.
 PA (OKUB/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 2027: 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 CC Sequence 337 BP; 104 A; 53 C; 59 G; 110 T;

Query Match 100.0%; Score 9; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 111 ttctgagaa 119
 |||||
 Qy 25 ttctgagaa 33

RESULT 69
 ID V24222 standard; DNA; 338 BP.
 AC V24222:
 DT 12-AUG-1998 (first entry)
 DE Fragment of centromere from human chromosome 21.
 KW CENP-B box; mammalian artificial chromosome; MAC; chromosome 21;
 KW mammalian telomere; centromere; yeast artificial chromosomes; YAC;
 KW somatic gene therapy; ds.
 OS Homo sapiens.
 PN W09808964-A1.
 PD 05-MAR-1998.
 PF 26-AUG-1996: J02381.
 PR 26-AUG-1996: WO-002381.
 PA (KENO/) IKENO M.
 PA (MASU/) MASUMOTO H.
 PA (OKAZ/) OKAZAKI T.
 PI Cooke HJ, Grimes BR, Ikono M, Masumoto H, Okazaki T;
 DR WPI: 98-216941/19.
 PT Mammalian artificial chromosomes - prepared using DNA construct
 PT comprising mammalian telomere and centromere containing copies of
 PT CENP-B box sequences
 PS Disclosure: Page 56; 95pp: English.
 CC This sequence is part of the human chromosome 21 centromere, and can be
 CC used in the DNA construct of the invention. The construct is a mammalian
 CC artificial chromosome (MAC), and comprises a mammalian telomere and a
 CC centromere, where the centromere has a DNA sequence containing copies of
 CC the CENP-B box sequence. The invention also relates to methods for
 CC establishing yeast artificial chromosomes (YACs) comprising mammalian
 CC telomeres and centromeres. The method can be used to construct, modify
 CC and stably maintain YACs in yeast cells which have the ability to form
 CC MACs when introduced into mammalian cells. The MACs can replicate
 CC autonomously, be stably maintained extrachromosomally and transmitted
 CC efficiently in mammalian cells. The MACs can be used for basic studies on
 CC organisation and function of mammalian chromosomes and also as a vector
 CC to introduce DNA segments (genes) of interest to test their functions in
 CC mammalian cells or bodies, since the genes carried by MACs will neither
 CC be subject to variable expression due to integration position effect nor
 CC cause unpredictable insertion mutation on host chromosomes. Furthermore
 CC MACs will have the capacity to accommodate a DNA segment up to megabases
 CC where an entire large gene or group of genes and regulatory elements
 CC could be included. The MACs can be used for somatic gene therapy or to
 CC generate transgenic mice.
 SQ Sequence 338 BP; 103 A; 58 C; 74 G; 103 T;

Query Match 100.0%; Score 9; DB 43; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 198 ttctgagaa 206
 |||||
 Cp 33 ttctgagaa 25

RESULT 70
ID V24223 standard; DNA: 339 BP.
AC V24223;
DE 12-AUG-1998 (first entry)
DE Fragment of centromere from human chromosome 21.
KW CENP-B box; mammalian artificial chromosome; MAC; chromosome 21;
KW mammalian telomere; centromere; yeast artificial chromosomes; YAC;
KW somatic gene therapy; ds.
OS Homo sapiens.
PN W09808964-A1.
PD 05-MAR-1998.
PR 26-AUG-1996: J02381.
PR 26-AUG-1996: W0-J02381.
PA (KNO/) IKENO M.
PA (MASU/) MASUMOTO H.
PA (OKAZ/) OKAZAKI T.
PI Cooke HJ, Grimes BR, Ikeno M, Masumoto H, Okazaki T;
PI WPI: 98-216941/19
DR Mammalian artificial chromosomes - prepared using DNA construct
PT Comprising mammalian telomere and centromere containing copies of
PT CENP-B box sequences
PS Disclosure: Page 57; 95pp: English.
CC This sequence is part of the human chromosome 21 centromere, and can be
CC used in the DNA construct of the invention. The construct is a mammalian
CC artificial chromosome (MAC), and comprises a mammalian telomere and a
CC centromere, where the centromere has a DNA sequence containing copies of
CC the CENP-B box sequence. The invention also relates to methods for
CC establishing yeast artificial chromosomes (YACs) comprising mammalian
CC telomeres and centromeres. The method can be used to construct, modify
CC and stably maintain YACs in yeast cells which have the ability to form
CC MACs when introduced into mammalian cells. The MACs can replicate
CC autonomously, be stably maintained extrachromosomally and transmitted
CC efficiently in mammalian cells. The MACs can be used for basic studies on
CC organization and function of mammalian chromosomes and also as a vector
CC to introduce DNA segments (genes) of interest to test their functions in
CC mammalian cells or bodies, since the genes carried by MACs will neither
CC be subject to variable expression due to integration position effect nor
CC cause unpredictable insertion mutation on host chromosomes. Furthermore
CC MACs will have the capacity to accommodate a DNA segment up to megabases
CC where an entire large gene or group of genes and regulatory elements
CC could be included. The MACs can be used for somatic gene therapy or to
CC generate transgenic mice.
SQ Sequence 339 BP: 103 A: 58 C: 73 G: 105 T:
Query Match 100.0%; Score 9; DB 43; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 199 ttctcagaa 207
| | | | | | | |
Cp 33 ttctcagaa 25

RESULT 71
ID T19986 standard; cDNA to mRNA: 339 BP.
AC T19986;
DE 17-JUL-1996 (first entry)
DE Human gene signature HUMGS01120.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994: J01916.
PR 12-NOV-1993: JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 529; 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 339 BP: 107 A: 63 C: 51 G: 116 T:
Query Match 100.0%; Score 9; DB 19; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 98 ttctcagaa 106
| | | | | | | |
Cp 33 ttctcagaa 25

RESULT 72
ID Q26400 standard; DNA: 342 BP.
AC Q26400;
DE 15-JAN-1993 (first entry)
DE GBSs gene fragment I from potato.
KW Granule-bound starch synthase; amylopectin; amylose production;
KW inhibition; ss.
OS Solanum tuberosum.
FH Key Location/Qualifiers
FT 217..342
FT m5c-feature
FT /tag- a
FT /product- GBSs_N-terminal-fragment
PN W09211376-A.
PD 09-JUL-1992.
PR 20-DEC-1991: SE0892.
PR 21-DEC-1990: SE-004096.
PA (AMYL-) AMYLOGENE HB.
PI Hofvander P, Persson PT, Tallberg A, Wikstrom O;
PI WPI: 92-250096/30.
DR P-PSDB: R25473.
PT Modifying potatoes to form amylopectin starch - using an
PT anti-sense construct to inhibit granule-bound starch synthase
PS Claim 1: Page 21; 46pp: English.
CC A genomic library in EMB3 was prepared using leaves of the potato
CC Bintje. The library was screened with cDNA clones for the 5' and 3'
CC ends of the GBSs gene. A full-length clone of potato GBSs gene was
CC identified (wx311). An EcoRI fragment ("w") of the clones was found to
CC contain the start of the gene and was cloned in pUC13 to give pSV.
CC Restriction of pSV with NsiI and HindIII gives fragment I which was
CC cloned in pUC19 to give 19NH35. Further restriction of 19NH35 with
CC HpaI-SstI gives a 342bp fragment comprising leader sequence, the
CC translation start and the first 125bp of the coding region. The
CC fragment can be used to make antisense constructs to suppress amylose
CC formation in potatoes. See also Q26401-4.
SQ Sequence 342 BP: 85 A: 78 C: 54 G: 125 T:
Query Match 100.0%; Score 9; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 132 ttctcagaa 140
| | | | | | | |
Cp 33 ttctcagaa 25

RESULT 73
ID Q60547 standard; cDNA; 343 BP.
AC Q60547;
DT 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST02554.
KW Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal; ss.
OS Homo sapiens.
PN WO9316178-A.
PD 19-AUG-1993.
PR 12-FEB-1993; U01294.
PT 12-FEB-1992; US-837195.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ;
DR WPI: 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
markers for human genes transcribed in-vivo, facilitate tagging
of most human genes.
PS Example 4; Page 351; 500pp; English.
CC The Expressed Sequence Tag was isolated from a human brain cDNA
library as part of a large set of ESTs which can be used as markers
for human genes transcribed in vivo. They can be used to facilitate
tagging of most human genes, for mapping locations of expressed genes
on chromosomes, for individual or forensic identification, for mapping
locations of disease-associated genes, for identification of tissue
type, and for prep. of antisense sequences, probes and constructs.
CC EST02554 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also Q59041-Q61440.
SQ Sequence 343 BP; 91 A; 70 C; 95 G; 84 T.

Query Match 100.0%; Score 9; DB 8; Length 343;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 233 tctcagaa 241
|||||
Oy 25 tctcagaa 33

RESULT 74
ID V69006 standard; DNA; 348 BP.
AC V69006;
DT 22-JAN-1999 (first entry)
DE DNA molecule encoding a polypeptide for detecting breast cancer #8.
KW Human; breast cancer; breast tumour tissue; diagnosis; treatment;
vaccine; epitope; endogenous; retroviral element; ss.
OS Homo sapiens.
PN MO9845328-A2.
PD 15-OCT-1998.
PR 09-APR-1998; U06939.
PR 11-DEC-1997; US-991789.
PR 09-APR-1997; US-838762.
PA (CORI-) CORIXA CORP.
PI Fridakis TN, Reed SG, Smith JM;
DR WPI: 98-557473/47.
PT New DNA sequences isolated from endogenous human retroviral element
PT - and related vectors, transformed cells, proteins and antibodies,
PT useful for diagnosis, treatment and prevention of breast cancer
PS Claim 11; Page 65; 173pp; English.
CC V68999 to V69038 represent nucleotide sequences which encode polypeptides
used in detecting human breast cancer. Detection or measurement of
human breast tumour specific polypeptides and nucleotide sequences,
or the corresponding RNA in a sample, is used for diagnosis and
CC monitoring of breast cancer. Human breast tumour specific polypeptides
CC and nucleotide sequences, and the vectors containing the DNAs, are also
CC useful in vaccines for inhibiting development (for prevention or
therapy) of breast cancer. The polypeptides may also be used to
CC raise monoclonal antibodies, used as immunoassay reagents.
SQ Sequence 348 BP; 105 A; 60 C; 74 G; 104 T;

Query Match 100.0%; Score 9; DB 52; Length 348;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 230 tctcagaa 238
|||||
Cp 33 tctcagaa 25

RESULT 75
ID V88650 standard; cDNA; 359 BP.
AC V88650;
DT 12-FEB-1999 (first entry)
DE EST clone G278.
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; chromatolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
OS Homo sapiens.
PN WO9845437-A2.
PD 15-OCT-1998.
PR 10-APR-1998; U06956.
PR 10-APR-1997; US-837312.
PA (GENT) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallee ER, McCoy JM, Merberg D,
PI Racine LA, Spaulding V, Treacy M;
DR WPI: 99-070078/06.
PT New polynucleotides encoding human secreted proteins - derived from
PT ovary, pituitary, retina and colon cDNA libraries
PS Claim 1; Page 467; 641pp; English.
CC The present sequence represents an expressed sequence tag (EST), and is
CC a polynucleotide of the invention. The polynucleotides of the invention
are all secreted EST sequences isolated from a variety of human tissue
sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
SQ Sequence 359 BP; 88 A; 85 C; 117 G; 67 T;

Query Match 100.0%; Score 9; DB 53; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 216 tctcagaa 224
|||||
Oy 25 tctcagaa 33

RESULT 76
ID Q59805 standard; cDNA; 361 BP.
AC Q59805;
DT 16-MAR-1994 (first entry)
DE Human brain Expressed Sequence Tag EST01547.
KW Gene transcription product; genetic markers; tagging; in vivo;
transcription; mapping; locations; chromosomes; chromosomal; ss.
OS Homo sapiens.
PN WO9316178-A.
PD 19-AUG-1993.
PR 12-FEB-1993; U01294.
PR 12-FEB-1992; US-837195.
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ;
DR WPI: 93-272882/34.
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
of most human genes

PS Example 4: Page 234: 500bp: English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes
 CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC See also 059041-061440.
 SQ Sequence 361 BP: 112 A: 74 C: 72 G: 101 T:
 Query Match 100.0%; Score 9: DB 8: Length 361;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 52 ttctcagaa 60
 |||||||||
 Cp 33 ttctcagaa 25
 RESULT 77
 ID T89417 standard: DNA: 371 BP.
 AC T89417:
 DT 30-MAR-1998 (first entry)
 DE Human chemotactic cytokine III (CCIII) encoding DNA.
 KM Chemotactic cytokine III: CCIII: human: treatment: tumour; infection;
 KW T-cell mediated autoimmune disease; leukaemia; psoriasis; asthma; ss.
 OS Homo sapiens.
 FH Key
 FT CDS Location/Qualifiers
 FT 58..303
 FT sig_peptide /tag- a
 FT 58..141
 FT /note- b
 FT /note- "putative leader sequence"
 FT 142..300
 FT mal_peptide /tag- c
 FT WC09732993-A1.
 PN 12-SEP-1997.
 PF 05-MAR-1996: U02985.
 PR 05-MAR-1996: WO-U02985.
 PA (HUMA) HUMAN GENOME SCI INC.
 PI Gentz RL, Ni J, Yu G;
 DR WPI: 97-457539/42.
 DR P-PSDB: W31512.
 PT New nucleic acid encoding human chemotactic cytokine III - used for
 PT treatment of e.g. tumours, infections, autoimmune diseases etc.
 PS Claim 4: Pages 58-59, 73pp: English.
 CC This DNA encodes the human chemotactic cytokine III (CCIII). The CCIII
 CC attracts and activates macrophages (or their precursors), neutrophils,
 CC basophils and some lymphocytes. This DNA is used to produce
 CC recombinant CCIII (including in vivo production) for treatment of
 CC conditions associated with CCIII like tumours, chronic or parasitic
 CC infections, leukaemia, T-cell mediated autoimmune disease, psoriasis,
 CC asthma, allergy. Agonists of CCIII are used to regulate haematopoiesis,
 CC stimulate growth factor activity, inhibit angiogenesis or promote wound
 CC healing. Antagonists of CCIII are used where CCIII activity is excessive,
 CC for example, in glomerulonephritis, inflammation, cerebral ischaemia,
 CC HIV-1 related disease, arthritis, infections, autoimmune disease,
 CC endotoxic shock, atherosclerosis, allergy and bone marrow failure. Such
 CC diseases or susceptibility to them, can also be diagnosed by detecting
 CC mutations in the gene encoding CCIII, or by detecting abnormal levels of
 CC CCIII, e.g. using CCIII specific antibodies in usual immunoassays. The
 CC antibodies can also be used to identify or detect CCIII expressing clones
 CC and to purify CCIII. This encoding DNA is also useful for chromosome
 CC identification and for cloning the genomic sequence.
 SQ Sequence 371 BP: 67 A: 110 C: 112 G: 82 T:
 Query Match 100.0%; Score 9: DB 37: Length 371;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 177 ttctcagaa 185
 |||||||||

Cp 33 ttctcagaa 25
 RESULT 78
 ID V89589 standard: cDNA: 374 BP.
 AC V89589:
 DT 15-FEB-1999 (first entry)
 DE EST clone CN922.
 KW Human: secreted protein: expressed sequence tag: EST: haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 OS Homo sapiens.
 PN MO9845436-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U06955.
 PR 10-APR-1997: US-838821.
 PA (GEM) GENETICS INST INC.
 PI Agostino MJ, Jacobs R, Lavallie ER, McCoy JM, Merberg D,
 PI Racie LA, Spaulding V, Treacy M;
 DR WPI: 99-070077/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 PS Claim 1: Page 266: 618pp: English.
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 SQ Sequence 374 BP: 100 A: 85 G: 118 T:
 Query Match 100.0%; Score 9: DB 53: Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 231 ttctcagaa 239
 |||||||||
 Cp 33 ttctcagaa 25
 RESULT 79
 ID V36486 standard: cDNA: 375 BP.
 AC V36486:
 DT 28-SEP-1998 (first entry)
 DE Partial nucleosomal sequence from CP140 partial cDNA.
 KM Partial nucleosomal protein 140: CP140: cloning; screening; scleroderma;
 KW Bluescript vector; anti-CP140 autoantibody; sclerotic disease; ss.
 OS Homo sapiens.
 FH Key
 FT CDS Location/Qualifiers
 FT 2..373
 FT /tag- a
 FT /product- "CP140 partial protein"
 FT /note- "no start or stop codons at 5' and 3' end of
 FT sequence"
 FT WC09820025-A1.
 PN 14-MAY-1998.
 PD 04-NOV-1997: U20520.
 PF 05-NOV-1996: US-743200.
 PR (UYMA-) UNIV MASSACHUSETTS.
 PA Doxsey SJ;
 PI Doxsey SJ;
 DR P-PSDB: W61175.
 PT Screening for sclerotic disease in patients - by detection of
 PT anti-CP140 autoantibodies, altered CP140 mRNA, altered CP140 protein.
 PT levels or an alteration in the CP140 gene

PS Disclosure: Page 28-29: 54pp: English.
 CC The nucleic acids V36481-V36487 encode a variety of partial
 CC polynucleotides from the partial centrosomal protein 140 (CP140) cDNA.
 CC The CP140 partial cDNA was found by standard cloning and screening
 CC techniques. This sequence was then subcloned into a Bluescript vector.
 CC The 1.7 kb cDNA was sequenced and analysed and it was found that this
 CC partial sequence contained one continuous open reading frame. The
 CC polynucleotides and polypeptides (W61170-W61176) can be used as a
 CC diagnostic method for screening a patient for the presence of anti-CP140
 CC autoantibodies as an indication of sclerotic disease. This can be
 CC performed by taking a sample from the patient, adding to this a pure
 CC CP140 fragment, and then seeing if any antibodies present will bind the
 CC CP140 fragment. If any complexes are detected, it indicates that the
 CC patient is probably suffering from sclerotic disease.
 SQ Sequence 375 BP: 131 A: 83 C: 88 G: 73 T;

Query Match 100.0%; Score 9; DB 45; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 236 tctctcagaa 244
 Cp 33 tctctcagaa 25

RESULT 80
 ID T67613 standard; DNA: 381 BP.
 AC T67613:
 DT 14-JUL-1997 (first entry)
 DE H. pylori cytoplasmic protein ORF 34574062.aa.
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; genome;
 KW replication; transcription; recombination; repair; ds.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT cds 1..381
 FT /tag= a
 FT /transl_except= (pos.: 361..363, a.a.: Ser)
 FT /note= "No stop codon given in sequence"
 FT 364..366
 FT /tag= b
 FT /note= "encodes Arg"

PN W09640893-A1.
 PD 19-DEC-1996
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASPR) ASPRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL;
 DR WPI: 97-052306/05.
 DR P-PSDB: W20440.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 9: Pages 265: 1481pp: English.
 CC The present sequence encodes a Helicobacter pylori cytoplasmic protein
 CC involved in genome replication, transcription, recombination and repair.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, useful
 CC as potential H. pylori life cycle activators or inhibitors. The genomic
 CC sequence of H. pylori (ATCC 55679) was determined from overlapping
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 381 BP: 96 A: 86 C: 91 G: 106 T;

Query Match 100.0%; Score 9; DB 30; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 tctctcagaa 293
 Cp 33 tctctcagaa 25

RESULT 81
 ID V87945 standard; cDNA: 389 BP.
 AC V87945:
 DT 12-FEB-1999 (first entry)
 DE EST clone FG965.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; actlytin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN W09845437-A2.
 PD 15-OCT-1998
 PF 10-APR-1998; U06956.
 PR 10-APR-1997; US-837312.
 PA (GENY) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 DR Racie LA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1; Page 231; 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC actlytin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.

CC Sequence 389 BP: 112 A: 83 C: 93 G: 101 T;

Query Match 100.0%; Score 9; DB 54; Length 389;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 370 tctctcagaa 378
 Cp 25 tctctcagaa 33

RESULT 82
 ID V88795 standard; cDNA: 412 BP.
 AC V88795:
 DT 12-FEB-1999 (first entry)
 DE EST clone H0640.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; actlytin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN W09845437-A2.
 PD 15-OCT-1998
 PF 10-APR-1998; U06956.
 PR 10-APR-1997; US-837312.
 PA (GENY) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 DR Racie LA, Spaulding V, Treacy M;

DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain.
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1: Page 515; 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SQ Sequence 412 BP; 110 A; 71 C; 113 G; 118 T;
 Query Match 100.0%; Score 9; DB 53; Length 412;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 367 ttctgagaa 375
 Oy 25 ttctgagaa 33
 RESULT 83
 ID T19076 standard; cDNA to mRNA; 413 BP.
 AC T19076.
 DT 20-JUN-1996 (first entry)
 DE Human gene signature H0MG500088.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN M09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATSU) MATSUBARA K.
 PA (OKUBA) OKUBO K.
 PI Matsubara K; Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 293; 2245pp: Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 413 BP; 99 A; 64 C; 75 G; 154 T;
 Query Match 100.0%; Score 9; DB 18; Length 413;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 ttctgagaa 24
 Oy 25 ttctgagaa 33
 RESULT 84
 ID T15982 standard; DNA; 415 BP.
 AC T15982.
 DT 25-SEP-1996 (first entry)
 DE PSKAN8 fragment #1.
 KW Human; hPSTI; pancreatic secretory trypsin inhibitor; pIII protein;
 KW protease; vaccine; trypsin-inhibitory loop; alpha-chymotrypsin; elastase;
 KW antigen; M13; ss.
 OS Synthetic.
 FH Key location/Qualifiers
 FT primer_bind 15..37
 FT /tag- a
 FT /note- "binding site for primer #533 (see T15973)"
 FT cds 112..415
 FT /tag- b
 FT /product- PSKAN8 fusion protein fragment
 FT misc_feature 112..174
 FT /tag- c
 FT /note- "coding sequence for ompa-leader"
 FT misc_feature 175..342
 FT /tag- d
 FT /note- "coding sequence for pstI"
 FT primer_bind 187..218
 FT /tag- e
 FT /note- "binding site for primer #528 (see T15972)"
 FT primer_bind 200..222
 FT /tag- f
 FT /note- "binding site for primer #529 (see T15983)"
 FT primer_bind 335..370
 FT /tag- g
 FT /note- "binding site for mutagenic primer #1243"
 FT misc_feature 343..393
 FT /tag- h
 FT /note- "linker"
 FT misc_feature 394..415
 FT /tag- i
 FT /note- "P1II coding sequence"
 FT primer_bind 407..414
 FT /tag- j
 FT /note- "binding site for primer #1255 (see T15974)"
 FT misc_feature 415
 FT /tag- k
 FT /note- "1285 unspecified bases to join to T35685 to
 FT create full length pIII"
 FT EP-699750-A1.
 PN 06-MAR-1996.
 PD 07-JUN-1995; 108761.
 PR 07-JUN-1994; EP-108689.
 PA (GBFB) GBF GES BIOTECHE FORSCHUNG GMBH.
 PI Collins J; Roetgen P;
 DR WPI: 96-130768/14.
 P-PSB; R92296.
 PT filamentous DNA bacteriophage protein - have defined target
 PT phagendous DNA bacteriophage protein - have defined target
 PT molecules such as protease(s), e.g. alpha-chymotrypsin or elastase
 PT Example; Fig 1c; 19pp: English.
 PS CC T15983 and T35685 represent fragments of the phagemid vector PSKAN8.
 CC PSKAN8 contains a fusion between the human pancreatic secretory trypsin
 CC inhibitor (hPSTI) gene and the M13 pIII protein gene. In order to create
 CC a randomised phagemid bank from PSKAN8, expression cassettes containing
 CC NKK repeats were created and inserted into PSKAN8. The expression
 CC cassettes were created by using the sequences represented by
 CC T15983-T15987 as amplification primers, and used to randomise the region
 CC coding the trypsin-inhibitory loop of PSKAN8. The products of these
 CC phagemids can be used to isolate phagemid particles with strong binding
 CC characteristics for defined target molecules, such as proteases (e.g.
 CC alpha-chymotrypsin or elastase). The phagemid particles can then be used
 CC to isolate protein ligands. The phagemid banks present a combined

CC mutational density and diversity in excess of that in similar banks, such
 CC as immunoglobulin presenting phagemids. Using the phagemids it is
 CC possible to produce antigenic components of vaccines in which only the
 CC mutated region is antigenic, the majority of the protein being a
 CC correctly folded human protein with post-translational modification
 CC sites.
 SO Sequence 415 BP; 124 A; 92 C; 102 G; 96 T;
 Query Match 100.0%; Score 9; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 364 ttctcagaa 372
 Cp 33 ttctcagaa 25
 RESULT 85
 ID V44633 standard; DNA: 420 BP.
 AC V44633:
 DT 24-NOV-1998 (first entry)
 DE Human uncoupling protein-2 gene clone hUCP2-g1 sequence pSUB23sp6.
 KW Uncoupling protein-2; UCP2 gene; human; respiration;
 KW thermogenesis; obesity; hyperinsulinaemia; glucose intolerance;
 KW diabetes; syndrome X; hypothermia; wasting; cachexia; anorexia;
 KW inflammation; fever; hyperthermia; gene therapy; diagnosis; ds.
 OS Homo sapiens.
 PN W09831396-A1.
 PD 23-JUL-1998.
 PR 22-APR-1997; U06864.
 PA 15-JAN-1997; US-034960.
 PA (NARE-) CENT NAT RECH SCI CENT RECH SUR ENDOCRINOL.
 PA (REGC-) UNIV CALIFORNIA.
 PA (UTDU-) UNIV DUKE.
 PI Bouillaud F, Collins SA, Riequier D, Seidln MF,
 PI Surwit RS, Warden CH;
 DR MPI: 98-413823/35.
 PT Method for treating disease associated with altered UCP-2 expression
 PT - by administering agent which enhances or inhibits UCP-2 activity,
 PT effectively to treat obesity, diabetes, fever, hyperthermia,
 PT cachexia etc.
 PS Disclosure: Fig 1d; 98pp; English.
 CC This is the nucleotide sequence of subcloned fragment pSUB23sp6 of
 CC the human uncoupling protein-2 (UCP2) gene present in genomic clone
 CC hUCP2-g1 (1-1806). 10 subcloned fragments (see V44633-38) of
 CC hUCP2-g1 are provided. hUCP2-g1 was isolated from a human placenta
 CC genomic DNA library constructed in EMBL3 phage using mouse UCP2
 CC cDNA as probe. pSUB23sp6 is a non-localised fragment of the gene.
 CC The human UCP2 gene maps to a chromosomal region (11q13) linked to
 CC obesity and hyperinsulinaemia. The invention provides methods for
 CC the treatment of disorders associated with diminished or elevated
 CC UCP2 expression or activity. An agent which enhances UCP2
 CC expression (e.g. an expression construct comprising a UCP2 encoding
 CC sequence) can be used to treat obesity, diabetes, syndrome X,
 CC hyperthermia, hyperinsulinaemia, or glucose intolerance. An
 CC inhibitor of UCP2 (e.g. an antisense construct) is used to treat
 CC wasting, anorexia, inflammation, cachexia, fever or hyperthermia
 CC (all claimed). The invention also relates to diagnostic and drug
 CC screening methodologies.
 SO Sequence 420 BP; 109 A; 95 C; 94 G; 121 T;
 Query Match 100.0%; Score 9; DB 48; Length 420;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 200 ttctcagaa 208
 Cp 33 ttctcagaa 25
 RESULT 86
 ID T64806 standard; DNA: 424 BP.
 AC T64806:

DT 03-SEP-1997 (first entry)
 DE Rat RC-9 DNA useful in proliferative arterial disease diagnosis.
 KW RC-9; proliferative arterial disease; vascular stenosis;
 KW diagnosis; therapy; ds.
 OS Rattus sp.
 PN W09722880-A1.
 PD 26-JUN-1997.
 PR 13-DEC-1996; U19671.
 PA 18-DEC-1995; US-008801.
 PA (SMK-) SMITHKLINE BEECHAM CORP.
 PI Arleth AJ, Autieri MV, Ohlstein EH;
 DR MPI: 97-341832/31.
 PT Diagnosis and treatment of proliferative arterial disease - using
 PT RC-9 polynucleotide(s), proteins and antibodies, e.g. to identify
 PT inhibitors of RC-9 activity
 PS Claim 4; Page 19-20; 34pp; English.
 CC RC-9 polynucleotides (T64806-09) correspond to partial, genomic
 CC and cDNA sequences from rat, and human cDNA, respectively. The
 CC rat 424 bp partial genomic DNA fragment was isolated by
 CC differential display analysis of rat carotid arteries after balloon
 CC angioplasty and was used as a probe to isolate the genomic
 CC sequence. RC-9 protein (see also W18513-14) is implicated in
 CC coronary arterial disease. RC-9 nucleic acids can be used in
 CC methods for the prodn. of recombinant RC-9. In the diagnosis and
 CC treatment of proliferative arterial disease, and in the
 CC identification of RC-9 (ant)agonists useful in therapeutic compns.
 SO Sequence 424 BP; 137 A; 80 C; 112 G; 95 T;
 Query Match 100.0%; Score 9; DB 31; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 169 ttctcagaa 177
 Cp 33 ttctcagaa 25
 RESULT 87
 ID V02896 standard; cDNA: 430 BP.
 AC V02896:
 DT 08-MAY-1998 (first entry)
 DE Human HMG1-C aberrant form 17.
 KW High mobility group protein; HMG1-C; HMG; human; treatment; modulator;
 KW multiple tumour aberration growth gene; vascular development;
 KW angiogenesis; vascularisation; endometriosis; contraception
 KW tissue regeneration; ss.
 OS Homo sapiens.
 PN D019548122-A1.
 PD 26-JUN-1997.
 PR 21-DEC-1995; 048122.
 PR 21-DEC-1995; DE-048122.
 PA (BULL/) BULLERDREK J.
 PI Bullerdiek J;
 DR MPI: 97-333837/31.
 PT DNA sequences representing aberrant forms of human high mobility
 PT group protein genes - useful for treatment of endometriosis and
 PT tumours, or for modulating vascularisation, etc
 PS Claim 1; Fig 17; 58pp; German.
 CC V02880-V02896 are cDNA sequences that encode aberrant forms of the human
 CC high mobility group protein (HMG) gene, HMG1-C, which is located on
 CC chromosome 12. These sequences encode the DNA binding part of the
 CC translation product but not the protein binding domain. These proteins,
 CC antibodies derived from these proteins or expression modulators of the
 CC HMG1-C protein can be used in kits to modulate vascular development. Such
 CC kits can reduce, block or stimulate angiogenesis or vascularisation and
 CC can improve vascular provision in myocardium damaged by infarction. Such
 CC proteins can also be used to treat endometriosis and tumours, for
 CC contraception (local or oral) and for tissue regeneration, especially in
 CC degenerating or damaged tissue. The regeneration method can be applied to
 CC tissues which are currently impossible or difficult to regenerate and the
 CC use of biological material with attendant risks of viral transmission and
 CC anaphylactic shock, is avoided.
 SO Sequence 430 BP; 129 A; 108 C; 113 G; 80 T;

Query Match 100.0%; Score 9; DB 38; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 380 ttctcagaa 388
|||||||
Cc 33 ttctcagaa 25

RESULT 88
ID V34161 standard; DNA: 436 bp.
AC V34161;
DT 28-JAN-1999 (first entry)
DE Human secreted protein gene 8 clone HMBX59.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN M09839446-A2.
PD 11-SEP-1998.
PF 06-MAR-1998; US-04492.
PR 07-MAR-1997; US-038621.
PR 07-MAR-1997; US-040161.
PR 07-MAR-1997; US-040162.
PR 07-MAR-1997; US-040163.
PR 07-MAR-1997; US-040333.
PR 07-MAR-1997; US-040334.
PR 07-MAR-1997; US-040336.
PR 07-MAR-1997; US-040626.
PR 11-APR-1997; US-043311.
PR 11-APR-1997; US-043312.
PR 11-APR-1997; US-043313.
PR 11-APR-1997; US-043314.
PR 11-APR-1997; US-043315.
PR 11-APR-1997; US-043568.
PR 11-APR-1997; US-043569.
PR 11-APR-1997; US-043576.
PR 11-APR-1997; US-043578.
PR 11-APR-1997; US-043580.
PR 11-APR-1997; US-043670.
PR 11-APR-1997; US-043679.
PR 11-APR-1997; US-043671.
PR 11-APR-1997; US-043672.
PR 11-APR-1997; US-043674.
PR 23-MAY-1997; US-047492.
PR 23-MAY-1997; US-047500.
PR 23-MAY-1997; US-047501.
PR 23-MAY-1997; US-047502.
PR 23-MAY-1997; US-047503.
PR 23-MAY-1997; US-047581.
PR 23-MAY-1997; US-047582.
PR 23-MAY-1997; US-047583.
PR 23-MAY-1997; US-047584.
PR 23-MAY-1997; US-047585.
PR 23-MAY-1997; US-047586.
PR 23-MAY-1997; US-047587.
PR 23-MAY-1997; US-047588.
PR 23-MAY-1997; US-047589.
PR 23-MAY-1997; US-047590.
PR 23-MAY-1997; US-047592.
PR 23-MAY-1997; US-047593.
PR 23-MAY-1997; US-047594.
PR 23-MAY-1997; US-047595.
PR 23-MAY-1997; US-047596.
PR 23-MAY-1997; US-047597.
PR 23-MAY-1997; US-047598.
PR 23-MAY-1997; US-047599.

PR 23-MAY-1997; US-047600.
PR 23-MAY-1997; US-047601.
PR 23-MAY-1997; US-047612.
PR 23-MAY-1997; US-047613.
PR 23-MAY-1997; US-047614.
PR 23-MAY-1997; US-047615.
PR 23-MAY-1997; US-047617.
PR 23-MAY-1997; US-047618.
PR 23-MAY-1997; US-047632.
PR 23-MAY-1997; US-047633.
PR 23-MAY-1997; US-048964.
PR 06-JUN-1997; US-048974.
PR 22-AUG-1997; US-056630.
PR 22-AUG-1997; US-056631.
PR 22-AUG-1997; US-056632.
PR 22-AUG-1997; US-056636.
PR 22-AUG-1997; US-056637.
PR 22-AUG-1997; US-056662.
PR 22-AUG-1997; US-056664.
PR 22-AUG-1997; US-056845.
PR 22-AUG-1997; US-056862.
PR 22-AUG-1997; US-056864.
PR 22-AUG-1997; US-056872.
PR 22-AUG-1997; US-056874.
PR 22-AUG-1997; US-056875.
PR 22-AUG-1997; US-056876.
PR 22-AUG-1997; US-056877.
PR 22-AUG-1997; US-056878.
PR 22-AUG-1997; US-056879.
PR 22-AUG-1997; US-056880.
PR 22-AUG-1997; US-056881.
PR 22-AUG-1997; US-056882.
PR 22-AUG-1997; US-056884.
PR 22-AUG-1997; US-056886.
PR 22-AUG-1997; US-056887.
PR 22-AUG-1997; US-056888.
PR 22-AUG-1997; US-056889.
PR 22-AUG-1997; US-056892.
PR 22-AUG-1997; US-056893.
PR 22-AUG-1997; US-056894.
PR 22-AUG-1997; US-056903.
PR 22-AUG-1997; US-056908.
PR 22-AUG-1997; US-056909.
PR 22-AUG-1997; US-056910.
PR 22-AUG-1997; US-056911.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057651.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Bednarik DP, Brewer JA, Carter KC, Duan R, Edner R, Endress GA,
PI Feng P, Ferrle AM, Fischer CL, Graves KA, Greene JM, Hu JS,
PI Ryab H, Latleir DM, Li Y, Moore PA, Nl J, Olsen HS, Rosen CA,
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
DR WPI: 98-609887/51.
DR P-PSDB: W75064.
PT New isolated human genes and the secreted polypeptides they encode
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PT Claim 1; Page 170; 447pp; English.
PS This sequence represents a nucleic acid molecule which encodes a secreted
PS human protein. The gene number, and the clone it is derived from, are
PS detailed in the descriptor line. The gene can be used to generate fusion
PS proteins by linking to the gene to a human immunoglobulin Fc portion
PS (e.g. V34145) for increasing the stability of the fused protein as
PS compared to the human protein only.
CC The invention relates to 70 novel genes and their fragments (nucleic acid
CC sequences: V34154-V34276; amino acid sequences W75057-W75179) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 70 polynucleotides, based on
CC which tissues they are most highly expressed in (see V34154 for described
CC uses).

SQ Sequence 436 BP; 154 A; 86 C; 78 G; 118 T;
 Query Match 100.0%; Score 9; DB 52; Length 436;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 194 tctcagaa 202
 |||||
 OY 25 tctcagaa 33

RESULT 89
 ID N80067 standard; DNA: 444 BP.
 AC N80067;
 DT 16-NOV-1990 (first entry)
 DE Equine gamma interferon encoding sequence.
 KW Equine gamma interferon polypeptide; antiviral agent; antitumour agent;
 KW antipressive agent; EGI; ss.
 OS Equus caballus.
 PN DE3642096-A.
 PD 16-JUN-1988.
 PF 10-DEC-1986; 642096.
 PR 10-DEC-1986; DE-642096.
 PA (BOEH) Boehringer Ingelheim.
 PI Hauptmann R, Himmler A, Swetly P;
 DR WPI: 88-168693/25.
 DR P-P5DB: P82922.
 PT Polypeptide cpds. with equine gamma interferon activity - are antiviral,
 PT antitumour and antipressive for use in horses.
 PS Claim 18; Page 9; 39pp; German.
 CC DNA was extracted from horse liver, digested and fragments isolated
 CC to be cloned in lambda phages. These were packaged in E. coli and
 CC tested with a human GI probe. A 4.6 kb BamHI fragment, encoding the
 CC entire gene was cloned in pUC3 to form pAH11. The introns were
 CC removed from the corresponding mRNA and the fragments spliced to give
 CC a product which can be converted to "copy DNA".
 CC This sequence has bp 1-60 deleted and an Met encoding codon added.
 CC The product is useful as antiviral, antitumour and immunosuppressive
 CC agent in horses.
 SO Sequence 444 BP; 140 A; 108 C; 83 G; 113 T;

Query Match 100.0%; Score 9; DB 1; Length 444;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 405 tctcagaa 413
 |||||
 Cp 33 tctcagaa 25

RESULT 90
 ID Q59535 standard; cDNA: 446 BP.
 AC Q59535;
 DT 16-MAR-1994 (first entry)
 DE Human brain Expressed Sequence Tag EST00528.
 KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcription; mapping; locations; chromosomes; chromosomal; ss.
 OS Homo sapiens.
 PN W09316178-A.
 PD 19-AUG-1993.
 PF 12-FEB-1993; U01294.
 PR 12-FEB-1992; US-837195.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Adams MD, Moreno RF, Venter CJ;
 DR WPI: 93-272882/34.
 PT Enriched oligonucleotides and corresp. sequences - used as
 PT markers for human genes transcribed in-vivo, facilitate tagging
 PT of most human genes
 PS Example 4; Page 191; 500pp; English.
 CC The Expressed Sequence Tag was isolated from a human brain cDNA
 CC library as part of a large set of ESTs which can be used as markers
 CC for human genes transcribed in vivo. They can be used to facilitate
 CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosomes, for individual or forensic identification, for mapping
 CC locations of disease-associated genes, for identification of tissue
 CC type, and for prep. of antisense sequences, probes and constructs.
 CC EST00528 has a "marginal" coding probability as evaluated using the
 CC coding-region prediction program CRM. See also Q59041-Q61440.
 SQ Sequence 446 BP; 131 A; 100 C; 109 G; 106 T;

Query Match 100.0%; Score 9; DB 8; Length 446;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 274 tctcagaa 282
 |||||
 OY 25 tctcagaa 33

RESULT 91
 ID T19056 standard; cDNA to mRNA: 450 BP.
 AC T19056;
 DT 20-JUN-1996 (first entry)
 DE Human gene signature HUMGS00068.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATS/) MATSUBARA K.
 PI (OKUBO/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 95-106931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function. By preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1; Page 286; 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SO Sequence 450 BP; 111 A; 86 C; 110 G; 139 T;

Query Match 100.0%; Score 9; DB 18; Length 450;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 281 tctcagaa 289
 |||||
 OY 25 tctcagaa 33

RESULT 92
 ID V88140 standard; cDNA: 469 BP.
 AC V88140;
 DT 12-FEB-1999 (first entry)
 DE EST clone DQ276.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

OS Homo sapiens.
 PN W09845437-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U069956.
 PR 10-APR-1997: US-837312.
 PA (GEMT) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1: Page 296: 641pp: English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SQ Sequence 469 BP: 172 A: 83 C: 110 G: 104 T;
 Query Match 100.0%; Score 9; DB 53; Length 469;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 ttctcagaa 84
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 93
 ID Q29870 standard: DNA: 481 BP.
 AC Q29870;
 DT 15-MAR-1993 (first entry)
 DE Pheromone receptor clone J8.
 KM Odorant receptor; insect; vertebrate; fish; mammal; neurotransmitter;
 KM hormone; G-protein; surface receptor; olfactory epithelium; PCR;
 KM Sprague-Dawley rat; amplify; primer; polymerase chain reaction;
 KM Multigene family; ligand binding domain; ss.
 OS Rattus rattus.
 FH Key
 FT cds Location/Qualifiers
 FT 2..481
 FT /*tag- a
 PN W09217585-A.
 PD 15-OCT-1992.
 PF 06-APR-1992: U02741.
 PR 05-APR-1991: US-681880.
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Axel R, Buck LB;
 DR WPI: 92-366257/44.
 DR P-PSDB: R27882.
 PT Nucleic acid encoding an odorant receptor - can be used to
 PT control insect populations or for detecting odours e.g. alcohol,
 PT explosives, natural gas etc.
 PS Claim 19: Fig 24: 195pp: English.
 CC The sequences given in Q29855-77 are odorant receptor clones derived
 CC from an insect, a vertebrate, a fish or a mammal. These clones form
 CC a family of neurotransmitters and hormone receptors which transduce
 CC intracellular signals by activation of specific G-proteins. Each
 CC of these receptors is a member of a superfamily of surface receptors
 CC which traverse the membrane seven times. These clones are only
 CC expressed in the olfactory epithelium. These clones were isolated
 CC using probes derived from RNA prepared from the olfactory epithelia
 CC of Sprague-Dawley rats. Isolated cDNAs were amplified using primers
 CC which correspond to transmembrane domain 2 and 7. PCR products of the

CC appropriate size were sequenced. The deduced protein
 CC sequences of these cDNAs defined a new multigene family which shared
 CC sequence and structural properties with the superfamily of
 CC neurotransmitter and hormone receptors which traverse the membrane
 CC seven times. This novel family, however exhibits features different
 CC from any other member of the superfamily identified so far. There is a
 CC striking divergence within the third, fourth and fifth transmembrane
 CC domains between the olfactory proteins. This divergence in the
 CC potential ligand binding domain is consistent with the idea that
 CC the family of molecules cloned is capable of associating with a large
 CC number of odorant of diverse molecular structure.
 SQ Sequence 481 BP: 85 A: 158 C: 92 G: 141 T;
 Query Match 100.0%; Score 9; DB 5; Length 481;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 300 ttctgagaa 308
 |||||||||
 Oy 25 ttctgagaa 33

RESULT 94
 ID V63193 standard: cDNA: 483 BP.
 AC V63193;
 DT 13-JAN-1999 (first entry)
 DE cDNA from clone eh61.1 which encodes a secreted protein.
 KM Secreted protein; immune stimulating; suppressing;
 KM haematopoiesis regulating activity; tissue growth activity; activin;
 KM inhibin activity; chemotactic; chemokinetic activity; haemostatic;
 KM thrombolytic activity; anti-inflammatory activity; cadherin;
 KM tumour invasion suppressor activity; tumour inhibition activity; ds.
 OS Homo sapiens.
 FH Key
 FT cds Location/Qualifiers
 FT 85..450
 FT /*tag- a
 PN W09844113-A1.
 PD 08-OCT-1998.
 PF 27-MAR-1998: U06176.
 PR 25-MAR-1998: US-823330.
 PR 28-MAR-1997: US-823330.
 PA (GEMT) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
 DR WPI: 98-542703/46.
 DR P-PSDB: M80409.
 PT New isolated polynucleotide(s) and secreted proteins - are obtained
 PT from human cDNA libraries prepared from adult testes, foetal brain,
 PT adult brain, adult blood and placenta
 PT Claim 25: Page 82: 124pp: English.
 CC The present sequence encodes a secreted protein. The nucleic acid
 CC sequence is isolated from a human adult blood cDNA library using
 CC probe V63204. The polypeptide may have biological activities such as
 CC e.g. nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity or other activities.
 SQ Sequence 483 BP: 156 A: 84 C: 101 G: 142 T;
 Query Match 100.0%; Score 9; DB 51; Length 483;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 200 ttctgagaa 208
 |||||||||
 Oy 25 ttctgagaa 33

RESULT 95
 ID V88491 standard: cDNA: 484 BP.
 AC V88491;

DR 12-FEB-1999 (first entry)
DE EST clone GS4.
KM Expressed sequence tag; secreted protein; haematopoiesis regulator;
KM tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KM chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KM receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
OS Homo sapiens.
PN WO9845437-A2.
PD 15-OCT-1998.
PF 10-APR-1998; U06956.
PR 10-APR-1997; US-837312.
PA (GENY) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavalie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 99-070078/06.
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries
PS Claim 1: Page 414: 641pp: English.
CC The present sequence represents an expressed sequence tag (EST), and is
CC a polynucleotide of the invention. The polynucleotides of the invention
CC are all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity, haemostatic
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
SQ Sequence 484 BP; 93 A; 117 C; 101 G; 173 T;
Query Match 100.0%; Score 9; DB 53; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 169 ttctcagaa 177
|||||
CP 33 ttctcagaa 25
RESULT 96
ID V88866 standard; cDNA: 484 BP.
AC V88866;
DR 12-FEB-1999 (first entry)
DE EST clone IAA3.
KM Expressed sequence tag; secreted protein; haematopoiesis regulator;
KM tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KM chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KM receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
OS Homo sapiens.
PN WO9845437-A2.
PD 15-OCT-1998.
PF 10-APR-1998; U06956.
PR 10-APR-1997; US-837312.
PA (GENY) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavalie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 99-070078/06.
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries
PS Claim 1: Page 538: 641pp: English.
CC The present sequence represents an expressed sequence tag (EST), and is
CC a polynucleotide of the invention. The polynucleotides of the invention
CC are all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities

CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity, haemostatic
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
SQ Sequence 484 BP; 152 A; 112 C; 87 G; 133 T;
Query Match 100.0%; Score 9; DB 53; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 264 ttctcagaa 272
|||||
CP 33 ttctcagaa 25
RESULT 97
ID 063681 standard; DNA: 495 BP.
AC 063681;
DR 03-FEB-1995 (first entry)
DE Partial cDNA encoding S112a - a TR-interacting protein.
KM nuclear thyroid hormone interacting proteins; TR; J11; J12;
KM partial sequence; S112a: diagnosis; transcription factors;
KM transcriptional coactivator; treatment; thyroid hormone receptor;
KM thyroid related disorders; modulation; nuclear hormone receptor;
KM fusion proteins; isolation; ds.
OS Chimeric Homo sapiens.
OS Chimeric Bacterial sp.
FH Key
FH misc-feature 352..354
FT /tag- a
FT /note- "stop codon within ORF"
FT misc-feature 376..378
FT /tag- b
FT /note- "stop codon within ORF"
FT misc-feature 406..408
FT /tag- c
FT /note- "stop codon within ORF"
PN WO9410338-A.
PD 11-MAY-1994.
PE 29-OCT-1993; U10443.
PR 30-OCT-1992; US-969136.
PA (GENO) GEN HOSPITAL CORP.
PI Lee JW, Moore DD;
PI WPI: 94-199808/24.
DR P-PSDB: R53533.
DE Nuclear hormone receptor interacting polypeptides, esp. thyroid
PT hormone-interacting proteins (TRs) - for identifying proteins
PT useful in treatment and diagnosis of thyroid related disorders by
PT inoculating thyroid hormone receptor activity
PS Disclosure: Page 42-43; 105pp: English.
CC A number of TR-interacting proteins have been isolated and their
CC cDNA's partially sequenced. S112a gene shows no homology to any
CC known gene and is shown in this sequence. Nearly all the fusion
CC cDNAs isolated showed very strong dependence on hormone activation.
CC The proteins can be used in an in vivo trap system for the isolation
CC of proteins which associate with any nuclear hormone receptor. The
CC proteins and Abs may be used to treat or diagnose thyroid disorders,
CC and to modulate thyroid hormone receptor activity.
SQ Sequence 495 BP; 152 A; 130 C; 98 G; 114 T;
Query Match 100.0%; Score 9; DB 11; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 410 ttctcagaa 418
|||||
CP 33 ttctcagaa 25
RESULT 98


```

ID V82582 standard: DNA: 495 BP.
AC V82582;
DE 09-FEB-1999 (first entry)
DE Thyroid hormone receptor-interacting protein S112a partial DNA.
KW Thyroid hormone receptor-interacting protein; S112a;
OS Homo sapiens.
PN US846711-A.
PD 08-DEC-1998.
PF 04-APR-1994: 222719.
PR 30-OCT-1992: US-969136.
PA (GENO.) GEN HOSPITAL CORP.
PI Lee JW, Moore DP.
PI WPI: 99-059040/05.
DR P-PSDB: W85091.
PT Screening assay for nuclear hormone receptor modulators - using
PT cells containing reporter gene construct
PS Disclosure: Fig 4; 69pp; English.
CC The present sequence encodes thyroid hormone receptor-interacting
CC protein S112a. The protein was identified using the method of the
CC invention. The method is used to determine if a test protein is capable
CC of interacting with a nuclear hormone receptor protein in a
CC ligand-dependent manner. The method is used especially to screen for
CC proteins that interact with thyroid hormone receptors in a
CC ligand-dependent or ligand-sensitive manner.
SO Sequence 495 BP; 152 A; 130 C; 98 G; 114 T;

Query Match 100.0%; Score 9; DB 54; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 410 ttctcagaa 418
|||||
Cp 33 ttctcagaa 25

RESULT 99
ID 024293 standard: DNA: 498 BP.
AC 024293;
DE 27-OCT-1992 (first entry)
DE Ovine gamma-interferon gene.
KW IFN; cytokine; Immuno-depression; vaccine adjuvants; ss.
OS Ovis aries.
FH Key Location/Qualifiers
FT signal_peptide 1..60 /tag- a
FT /note- "putative signal sequence"
FT mat_peptide 61..495 /tag- b
FT /note- "mature peptide"
FT cds 1..498 /tag- c
FT /note- "Ovine gamma interferon"
FT misc_feature 1..27 /tag- d
FT /note- "IFN-N PCR primer"
FT /tag- e
FT misc_feature 477..492 /tag- e
FT /note- "IFN-C PCR primer"
FT /tag- e
PN W09205255-A.
PD 02-APR-1992.
PE 12-SEP-1991: AU00419.
PR 13-SEP-1990: AU-002294.
PR 21-MAR-1991: AU-005175.
PA (CSIR.) COMMONWEALTH SCIENT. ORG.
PI Roche J, Sew HF, Wood PR.
PI WPI: 92-150483/18.
DR P-PSDB: R23661.
PT DNA encoding ovine cytokine(s) - used to prepare recombinant
PT -cytokine(s) for treatment of immuno-depression in sheep
PS Claim 2; Fig 1; 79pp; English.
CC The sequence is that of the ovine gamma-interferon (IFN) gene which can

```

```

CC be used in the prodn. of recombinant gamma-interferon, a cytokine-like
CC molecule. The ovine cytokines will be useful in the intensive livestock
CC industries such as live animal export trade, feed-lots and intensive
CC rearing industries, where animals are subjected to great environmental
CC challenge with infectious diseases, partic. respiratory infections, and
CC are more prone to immunodepressive effects. The cytokines may be used
CC for treatment or prophylaxis to maintain, stimulate or enhance immuno-
CC responsiveness. They may also be useful as natural adjuvants for
CC vaccines for sheep and cattle. See also 024294-024318.
SO Sequence 498 BP; 162 A; 95 C; 109 G; 132 T;

Query Match 100.0%; Score 9; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 154 ttctcagaa 162
|||||
Cp 33 ttctcagaa 25

RESULT 101
ID V69758 standard: DNA: 498 BP.
AC V69758;
DE 01-MAR-1999 (first entry)
DE Tobacco ovary-specific PLE36 gene promoter.
KW PLE36 gene; promoter; tobacco; transgenic plant; seedless fruit;
KW tomato; watermelon; cucumber; ds.
OS Nicotiana tabacum.
PN W09849888-A1.
PD 12-NOV-1998.
PE 06-MAY-1998: U09013.
PR 06-MAY-1997: US-045725.
PA (UNIV.) UNIV KANSAS STATE RES FOUND.
PI Li Y;
PI WPI: 99-034673/03.
DR A new construct to express phytohormones in developing fruit -
PT useful for, e.g. producing substantially seedless fruit from
PT transgenic plants
PS Disclosure: Page 35; 49pp; English.
CC This is the nucleotide sequence of the ovary-specific transcription
CC regulatory element from the PLE36 gene of tobacco. A claimed DNA
CC construct comprises either an isopentenyl transferase (see W81575)
CC or a tryptophan oxygenase (see W81575)-encoding sequence, operably
CC linked to an ovary or developing fruit-specific plant-expressible
CC promoter (see also W69755 and W69759). The construct is used to
CC stably integrate enzymes involved in cytokinin or auxin biosynthesis
CC into the plant genome to achieve a transgenic plant (preferably
CC tomato, cucumber or watermelon) producing seedless fruit in the
CC absence of pollination.
SO Sequence 498 BP; 168 A; 114 C; 113 G; 103 T;

Query Match 100.0%; Score 9; DB 56; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 244 ttctcagaa 252
|||||
Cp 33 ttctcagaa 25

RESULT 101
ID N50023 standard: DNA: 501 BP.
AC N50023;
DE 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN-416.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiProliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501 /tag- a
FT /tag- a
PN EP-163993-A.

```


PD 11-DEC-1985. 105750.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50022.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral.
PS Claim 28; Chart 2a, page 32; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thymbmdga 457
|:|:|:|:|
Oy 25 ttctgagaa 33

RESULT 102
ID N50027 standard; DNA; 501 BP.
AC N50027;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 444.
KM Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT /*tag- a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50026.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2e, page 36; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 112 A; 31 C; 67 G; 80 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thymbmdga 457
|:|:|:|:|
Oy 25 ttctgagaa 33

RESULT 103
ID N50028 standard; DNA; 501 BP.
AC N50028;
DT 04-SEP-1991 (first entry)

DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 445.
KM Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT /*tag- a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50027.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2f, page 37; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 114 A; 31 C; 68 G; 79 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thymbmdga 457
|:|:|:|:|
Oy 25 ttctgagaa 33

RESULT 104
ID N50024 standard; DNA; 501 BP.
AC N50024;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 417.
KM Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT /*tag- a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50023.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2b, page 33; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 110 A; 32 C; 66 G; 81 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 449 thytbmga 457
|:|:|:|:|
QY 25 tctctgagaa 33

RESULT 105
ID N50025 standard; DNA: 501 BP.
AC N50025:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 418.
KW Antiviral; cell growth regulator; immune system regulator;
KM antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
ET EP-163993-A. /*tag- a
PD 11-DEC-1985.
PF 17-MAY-1985: 105750.
PR 17-MAY-1984: GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50024.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2c, page 34; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFS of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP: 112 A; 30 C; 69 G; 85 T;

Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thytbmga 457
|:|:|:|:|
QY 25 tctctgagaa 33

RESULT 106
ID N50032 standard; DNA: 501 BP.
AC N50032:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 449.
KW Antiviral; cell growth regulator; immune system regulator;
KM antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
ET EP-163993-A. /*tag- a
PD 11-DEC-1985.
PF 17-MAY-1985: 105750.
PR 17-MAY-1984: GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50031.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2j, page 41; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFS of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they

CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP: 108 A; 30 C; 72 G; 79 T;

Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thytbmga 457
|:|:|:|:|
QY 25 tctctgagaa 33

RESULT 107
ID N50031 standard; DNA: 501 BP.
AC N50031:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 448.
KW Antiviral; cell growth regulator; immune system regulator;
KM antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
ET EP-163993-A. /*tag- a
PD 11-DEC-1985.
PF 17-MAY-1985: 105750.
PR 17-MAY-1984: GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
DR P-PSDB: P50030.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2l, page 40; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFS of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP: 110 A; 30 C; 69 G; 80 T;

Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thytbmga 457
|:|:~::~|:|
QY 25 tctctgagaa 33

RESULT 108
ID N50023 standard; DNA: 501 BP.
AC N50023:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 410.
KW Antiviral; cell growth regulator; immune system regulator;
KM antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
ET EP-163993-A. /*tag- a
PD 11-DEC-1985.
PF 17-MAY-1985: 105750.
PR 17-MAY-1984: GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;

DR WPI: 85-311944/50.
P-PSDB: P50022.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2a, page 32; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thylmdgda 457
1:1:1:1:1
Cp 33 tctcagaa 25
RESULT 109
ID N50024 standard; DNA: 501 BP.
AC N50024:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN γ 417.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT cds /*tag= a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
P-PSDB: P50023.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2b, page 33; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 110 A; 32 C; 66 G; 81 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thylmdgda 457
1:1:1:1:1
Cp 33 tctcagaa 25
RESULT 110
ID N50027 standard; DNA: 501 BP.
AC N50027:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN γ 444.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.

FH Key Location/Qualifiers
FT cds 1..501
FT cds /*tag= a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
P-PSDB: P50026.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2e, page 35; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 112 A; 31 C; 67 G; 80 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thylmdgda 457
1:1:1:1:1
Cp 33 tctcagaa 25
RESULT 111
ID N50025 standard; DNA: 501 BP.
AC N50025:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN γ 418.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT cds /*tag= a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI: 85-311944/50.
P-PSDB: P50024.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2c, page 34; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 112 A; 30 C; 69 G; 85 T;
Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 449 thylmdgda 457
1:1:1:1:1
Cp 33 tctcagaa 25

```

RESULT 112
ID N50026 standard: DNA: 501 BP.
AC N50026;
DE 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFN $\gamma$  430.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 1..501
FT /tag= a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY-1985: 105750.
PR 17-MAY-1984: GB-012564.
PA (SEAR) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG:
PI WPI: 85-311944/50.
DR P-PSDB: P50025.
PT New modified human beta interferon polypeptide(s) - prep. by
PT plasmid transformed bacteria, with improved antiviral.
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2d, page 35; 71pp: English.
CC Compared with interferon beta prep. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index: improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 BP; 108 A; 31 C; 70 G; 81 T;

Query Match 100.0%; Score 9; DB 3; Length 501;
Best Local Similarity 44.4%; Pred. No. 1.37e+03;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 449 thtbcagaa 457
|:::|:|
Cp 33 ttctcagaa 25

RESULT 113
ID T27596 standard: CDNA: 503 BP.
AC T27596;
DE 21-AUG-1996 (first entry)
DE Partial human transforming growth factor beta receptor type II gene.
KW TGF-beta receptor; anti-cancer; inhibition; cytoplasmic domain;
KW hepatoma; ds.
OS Homo sapiens.
PN J08051978-A.
PD 27-FEB-1996.
PF 09-AUG-1994: 187031.
PR 09-AUG-1994: JP-187031.
PA (SAKA) OTSUKA PHARM CO LTD.
PA WPI: 96-174562/18.
DR P-PSDB: R96418.
PT Transforming growth factor-beta-receptor type II gene - useful for
PT growth inhibition of cancer cells
PS Claim 2; Page 9; 12pp: Japanese.
CC T27596 is a CDNA sequence encoding the signal sequence and cytoplasmic
CC domain of human transforming growth factor-beta (TGF-beta) receptor
CC type II. The sequence was isolated from hepatoma cell line HepG2, and
CC used as a probe for the isolation of the full coding sequence. The
CC TGF-beta receptor is useful for inhibiting the growth of cancer cells.
CC The gene is useful for the mass prodn. of TGF-beta receptor type II
CC protein.
SQ Sequence 503 BP; 145 A; 124 C; 117 G; 117 T;

Query Match 100.0%; Score 9; DB 20; Length 503;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 462 ttctcagaa 470

```

```

Cp 33 ttctcagaa 25

RESULT 114
ID V02884 standard: CDNA: 506 BP.
AC V02884;
DE 08-MAY-1998 (first entry)
DE Human HMG1-C aberrant form 5.
KW High mobility group protein; HMG1-C; MAG; human; treatment; modulator;
KW multiple tumor aberration growth gene; vascular development;
KW angiogenesis; vascularisation; endometrios; contraception
KW tissue regeneration; ss.
OS Homo sapiens.
PN DE19548122-A1.
PD 26-JUN-1997.
PF 21-DEC-1995: 048122.
PR 21-DEC-1995: DE-048122.
PA (BULL/) BULLERDIEK J.
PI Bullerdiex J;
DR WPI: 97-333637/31.
PT DNA sequences representing aberrant forms of human high mobility
PT group protein genes - useful for treatment of endometrios and
PT tumours, or for modulating vascularisation, etc
PS Claim 1; Fig 5; 58pp; Getman.
CC V02880-V02888 are CDNA sequences that encode aberrant forms of the human
CC high mobility group protein (HMG) gene, HMG1-C, which is located on
CC chromosome 12. These sequences encode the DNA binding part of the
CC translation product but not the protein binding domain. These proteins,
CC antibodies derived from these proteins or expression modulators of the
CC HMG1-C protein can be used in kits to modulate vascular development. Such
CC kits can reduce, block or stimulate angiogenesis or vascularisation and
CC can improve vascular provision in myocardium damaged by infarction. Such
CC proteins can also be used to treat endometrios and tumours, for
CC contraception (local or oral) and for tissue regeneration, especially in
CC degenerating or damaged tissue. The regeneration method can be applied to
CC tissues which are currently impossible or difficult to regenerate and the
CC use of biological material with attendant risks of viral transmission and
CC anaphylactic shock, is avoided.
SQ Sequence 506 BP; 167 A; 117 C; 124 G; 98 T;

Query Match 100.0%; Score 9; DB 38; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 357 ttctcagaa 365
|:::|:|
Cp 33 ttctcagaa 25

RESULT 115
ID T19874 standard: CDNA to mRNA: 514 BP.
AC T19874;
DE 26-JUL-1996 (first entry)
DE Human gene signature HDMS00997.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994: J01916.
PR 12-NOV-1993: JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUBA/) OKUBO K.
PI Matsubara K, Okubo K:
PI WPI: 95-206931/27.
DR Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 493; 2245pp: Japanese.
A single-stranded DNA (or its complementary strand or the corresp.

```

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types
 SO Sequence 514 BP; 136 A; 87 C; 92 G; 175 T;
 Query Match 100.0%; Score 9; DB 19; Length 514;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 33 tctcagaa 41
 Cp 33 tctcagaa 25
 RESULT 116
 ID X13325 standard; DNA: 519 BP.
 AC X13325;
 DT 19-MAR-1999 (first entry)
 DE Enterococcus faecalis genome contig SEQ ID NO:388.
 KW Enterococcus faecalis; contig; detection; Enterococcal Infection;
 KW vaccine; attenuation; computer readable medium; ds.
 OS Enterococcus faecalis.
 PN W08850555-A2.
 PD 12-NOV-1998;
 PF 04-MAY-1998; U088985.
 PR 14-NOV-1997; US-066009.
 PR 06-MAY-1997; US-044031.
 PR 16-MAY-1997; US-046655.
 PA (HUMAN) HUMAN GENOME SCI INC.
 PI Barash SC, Dillon PJ, Kunsch CA;
 DR WPI: 99-045171/04.
 PR New isolated Enterococcus faecalis polynucleotides and polypeptides
 PT - used to develop products for the detection of Enterococcus and for
 PT use in vaccines for prevention or attenuation of Enterococcus
 PT infection.
 PS Claim 1; Page 1571; 2084pp; English.
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
 CC X12938 to X13919 represent these nucleotide sequences which are primary
 CC nucleotide sequences, also known as contigs. The computer-based system
 CC can identify fragments of the Enterococcus faecalis genome with
 CC commercial importance. The products can be used to detect the presence
 CC of Enterococcus faecalis in samples. They can also be used for
 CC diagnosing Enterococcal infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 SO Sequence 519 BP; 146 A; 105 C; 83 G; 182 T;
 Query Match 100.0%; Score 9; DB 60; Length 519;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 391 tctcagaa 399
 Cp 25 tctcagaa 33
 RESULT 117

ID 045770 standard; DNA: 521 BP.
 AC 045770;
 DT 08-DEC-1993 (first entry)
 DE Human prostate transglutaminase gene clone PTC6561/2.
 KW Degenerative; cellular apoptosis; enzyme; inter alia;
 KW therapeutic wound repair; skin graft closure; food prepn.;
 KW preparation; stabilising; marker; identifying agent; agonists;
 KW antagonists; ss.
 OS Homo sapiens.
 PN W09313207-A.
 PD 08-JUL-1993.
 PF 30-DEC-1992; U11953.
 PR 31-DEC-1991; US-816284.
 PA (Zymo) ZYMOGENETICS INC.
 PI Grant RJ, Ohara PJ, Sheppard PO;
 DR WPI: 93-227323/28.
 PT Isolated polynucleotide molecule, for stabilising good prepn.
 PT utilised for coding human prostatic or placental trans glutaminase(s)
 PT and DNA, for repairing wounds, ulcerated lesions, skin grafts, and
 PT cellular markers
 PS Example: Page 36; 48pp; English.
 CC The sequence is that of the clone PTC6561/2 which contains a unique
 CC transglutaminase sequence and was isolated from a lambda-gt11 human
 CC prostate tissue cDNA library using primers ZC4127 and ZC4129. This
 CC was done as part of the generation of a full-length human prostate
 CC transglutaminase cDNA clone.
 SO Sequence 521 BP; 136 A; 125 C; 155 G; 105 T;
 Query Match 100.0%; Score 9; DB 7; Length 521;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 349 tctcagaa 357
 Cp 33 tctcagaa 25
 RESULT 118
 ID V87746 standard; cDNA: 534 BP.
 AC V87746;
 DT 12-FEB-1999 (first entry)
 DE EST clone EP234.
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; actinin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolytics;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PN W09845437-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998; U069956.
 PR 10-APR-1997; US-837312.
 PA (GENY) GENETICS INST INC.
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racie JA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PR New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1; Page 165; 64pp; English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.

SO Sequence 534 BP; 138 A; 148 C; 123 G; 125 T;
 Query Match 100.0%; Score 9; DB 54; Length 534;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 115 ttccagaa 123
 |||||||||
 Cp 33 ttccagaa 25

RESULT 119
 ID N60632 standard; DNA: 539 BP.
 AC N60632:
 DT 17-JUN-1991 (first entry)
 DE Beta-urogastrone - beta-lactamase fusion gene from pUG2701.
 KW Beta-urogastrone gene; hormone; salivary gland; ulcers; wounds;
 KM fusion protein; beta-lactamase; ss.
 OS Synthetic.
 FH Key
 FT cds
 FT 1.477
 FT /tag- a
 FT /product- fusion_protein
 FT cds
 FT 1.307
 FT /tag- b
 FT /product- beta-lactamase
 FT misc_rna
 FT 308..318
 FT /tag- c
 FT /label- adaptor
 FT cds
 FT 319..477
 FT /tag- d
 FT /product- beta-urogastrone

DE3523634-A.
 PD 09-JAN-1986; 523634.
 PF 02-JUL-1985; JP-137691.
 PR (EXRT-) EARTH CHEMICAL CO.
 PA Aoki S, Ohgai H, Horinaka A, Hiramatsu H, Kounoto S, Nishimura A,
 PI Matsushiro S;
 DR WPI: 86-015031/03.
 DR P-PSDB: P60628.
 DR New gene for expression of beta-urogastrone - its derivs., plasmid(s)
 PT and transformed cells contg. it.
 PS Disclosure: Page 59-61; 92pp; German.
 CC This sequence includes the sequence of claim 1 (bases 319-477)
 CC modified by restriction sites for coupling to the beta-lactamase
 CC gene and insertion into plasmids. Beta-urogastrone can then
 CC easily recovered from the fusion protein expressed by transformants.
 CC The fusion protein is less easily degraded by proteases and so
 CC protects beta-urogastrone and beta-lactamase collects in the periplasm
 CC of E.coli. It is therefore easy to collect and purify the product.
 CC Beta-urogastrone is the hormone of the salivary glands which suppresses
 CC stomach acid secretion and promotes cell growth, so is useful for
 CC treating ulcers and wounds. Previously the product was obtd. only
 CC in small amts. from human urine.
 CC See also N60628, and 30-32.
 SO Sequence 539 BP; 130 A; 116 C; 137 G; 156 T;

Query Match 100.0%; Score 9; DB 2; Length 539;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttccagaa 293
 |||||||||
 Cp 33 ttccagaa 25

RESULT 120
 ID T98787 standard; DNA: 541 BP.
 AC T98787:
 DT 10-NOV-1998 (first entry)
 DE DNA encoding a S. pneumoniae protein of unknown function.
 KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;

KW Immunological response; inoculation; antibody production; inhibitor;
 KW T cell immune response; antimicrobial compound; bacterial adhesion;
 KW extracellular matrix protein; protein-mediated cell invasion; wound;
 KW pathogenesis; ss.
 OS Streptococcus pneumoniae.
 FH Key
 FT cds
 FT 41..433
 FT /tag- a

NO9743303-A1.
 PN 20-NOV-1997.
 PD 14-MAY-1997; U07950.
 PR 14-MAY-1996; US-01670.
 PA (SMIR) SMITHKLINE BEECHAM CORP.
 PI (SMIR) SMITHKLINE BEECHAM PLC.
 PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,
 PI Stodola RK;
 DR WPI: 98-008793/01.
 DR P-PSDB: W38745.
 DR Novel Streptococcus pneumoniae proteins and related DNA - useful for
 PT diagnosing anti-microbial agents for treatment of bacterial
 PT infections

PS Claim 4; Page 258; 483pp; English.
 CC This sequence encodes a Streptococcus pneumoniae protein of unknown
 CC function, and represents a DNA sequence of the invention.
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain
 CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
 CC invention can be used to identify compounds which interact with and
 CC inhibit or activate the activity of the proteins. Antagonists can be
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic
 CC immunisation. They can also be used to induce an immunological response
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery
 CC of the encoding nucleic acids in a vector adequate to produce antibody
 CC and/or T cell immune responses to protect the animal from disease. The
 CC proteins can also be used to identify antimicrobial compounds which are
 CC capable of inhibiting their bioactivity. In particular the proteins of
 CC the invention can be used to prevent adhesion of bacteria to mammalian
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to
 CC block protein-mediated mammalian cell invasion, and to block the normal
 CC progression of pathogenesis in infections initiated other than by the
 CC implantation of in-dwelling devices or other surgical techniques.
 SO Sequence 541 BP; 136 A; 140 C; 91 G; 174 T;

Query Match 100.0%; Score 9; DB 48; Length 541;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 ttccagaa 227
 |||||||||
 Cp 33 ttccagaa 25

RESULT 121
 ID V68888 standard; DNA: 547 BP.
 AC V68888:
 DT 22-JAN-1999 (first entry)
 DE DNA molecule encoding a breast tumour specific polypeptide #80.
 KW Human; breast cancer; breast tumour tissue; diagnosis; treatment;
 KW vaccine; epitope; endogenous; retroviral element; ss.
 OS Homo sapiens.
 PN WO9845328-A2.
 PD 15-OCT-1998.
 PF 09-APR-1998; U06939.
 PR 11-DEC-1997; US-991789.
 PR 09-APR-1997; US-838762.
 PA (CORI-) CORIXA CORP.
 PI Frudakis TN, Reed SG, Smith JM;
 DR WPI: 98-557473/47.
 DR New DNA sequences isolated from endogenous human retroviral element
 PT - and related vectors, transformed cells, proteins and antibodies.
 PT useful for diagnosis, treatment and prevention of breast cancer
 PS Claim 1; Page 87; 173pp; English.
 DE V68880 to V68998 represent nucleotide sequences which encode human
 CC breast tumour specific polypeptides. Detection or measurement of

CC human breast tumour specific polypeptides and nucleotide sequences,
CC or the corresponding RNA in a sample, is used for diagnosis and
CC monitoring of breast cancer. Human breast tumour specific polypeptides
CC and nucleotide sequences, and the vectors containing the DNAs, are also
CC useful in vaccines for inhibiting development (for prevention or
CC therapy) of breast cancer. The polypeptides may also be used to
CC raise monoclonal antibodies, used as immunoassay reagents.
SO Sequence 547 BP; 154 A; 142 C; 91 G; 159 T;

Query Match 100.0%; Score 9; DB 52; Length 547;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 473 ttctcagaa 481
Cp 33 ttctcagaa 25

RESULT 122
ID Q22834 standard; cDNA; 556 BP.
AC Q22834:
DT 22-JUL-1992 (first entry)
DE Sequence encoding ovine interferon IFN gamma cDNA obtained by PCR.
KW Vaccine; antigen; therapeutic agent; immune response enhancer;
KW modulator; ss.
OS Ovis ammon aries.
FH Key Location/Qualifiers
FT cds 35..532
FT P-PSDB: R22133.
PD MO9203574-A.
PE 05-MAR-1992.
PE 13-AUG-1991: AU0358.
PR 13-AUG-1990: AU-001698.
PR 11-DEC-1980: AU-003859.
PR 15-FEB-1991: AU-004621.
PR 21-JUN-1991: AU-006840.
PA (UYN-) UNIV MELBOURNE.
PA (AMCO) AUSTRALIAN WOOL CORP.
PI Brandon MR, Andrews AE, Nash AD, Neusen EN;
DR WPI: 92-096916/12.
DR P-PSDB: R22133.
PT Nucleotide sequences coding for ruminant cytokine(s) or receptors
PT - used for producing polypeptide(s) for therapeutic and/or
PT adjuvant uses in animals
PS Disclosure: Fig 5D: 93pp: English.
CC The inventors claim a DNA sequence coding for a polypeptide
CC exhibiting ruminant cytokine or cytokine receptor activity; ovine
CC interleukine (IL)-1alpha activity; ovine IL-1alpha activity; ovine
CC IL-6 activity; ovine tumour necrosis factor (TNF) alpha activity;
CC ovine IL-2 receptor alpha activity; ovine interferon (IFN)-lambda
CC activity; or ovine IL-2 activity or homologous sequences, derivs. or
CC mutants, or fragments. The recombinant polypeptides are also
CC claimed.
SO Sequence 556 BP; 174 A; 114 C; 116 G; 152 T;

Query Match 100.0%; Score 9; DB 3; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 188 ttctcagaa 196
Cp 33 ttctcagaa 25

RESULT 123
ID X13601 standard; DNA; 567 BP.
AC X13601:
DT 19-MAR-1999 (first entry)
DE Enterococcus faecalis genome contig SEQ ID NO:664.
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
OS Enterococcus faecalis.
PN MO9850555-A2.

PD 12-NOV-1998.
PF 04-MAY-1998: U08985.
PR 14-NOV-1997: US-066009.
PR 06-MAY-1997: US-044031.
PR 16-MAY-1997: US-046653.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
DR WPI: 99-045171/04.
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
PS Claim 1: Page 1909-1910: 2084pp: English.
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC X12938 to X13919 represent these nucleotide sequences which are primary
CC nucleotide sequences, also known as contigs. The computer-based system
CC can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
SO Sequence 567 BP; 189 A; 95 C; 98 G; 182 T;

Query Match 100.0%; Score 9; DB 60; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 ttctcagaa 11
Cp 33 ttctcagaa 25

RESULT 124
ID V65298 standard; DNA; 582 BP.
AC V65298:
DT 24-DEC-1998 (first entry)
DE DNA encoding a S. pneumoniae protein of unknown function.
KW Streptococcus pneumoniae protein; recombinant; gene expression;
KW DNA chip; virulence; antibody; infection; detection; treatment; ss.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT cds 1..582
FT P-PSDB: W80616.
PD MO9826072-A1.
PD 18-JUN-1998.
PD 09-DEC-1997: U22578.
PR 13-DEC-1996: US-036281.
PA (EUIL) LILLY & CO ELI.
PI Baltz RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SR,
PI Mills BJ, Norris FH, Peery RB, Rokey PK, Rostek PR,
PI Skettrud PL, Smith MC, Solenberg PJ, Treadway PJ,
PI Young Bellido ML;
DR WPI: 98-348529/30.
DR P-PSDB: W80616.
PT Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT for evaluating gene expression, and identification of virulence
PT genes
PS Claim 2: Pages 187-188: 33pp: English.
CC This sequence encodes a Streptococcus pneumoniae protein of unknown
CC function. The invention provides DNA sequences (V65201 to V65304) from
CC the Streptococcus pneumoniae genome and corresponding protein sequences
CC (W60605 to W60728). A recombinant host containing a vector expressing any
CC of the above nucleic acids can be used for the recombinant expression of
CC the protein sequences. The invention also provides a DNA chip having
CC arrayed on it at least 15 base pair fragment of any one or more of these
CC DNA sequences. The DNA chip can be used methods for evaluating gene
CC expression in S. pneumoniae and for identifying virulence genes in

CC S. pneumoniae. Antibodies that selectively bind to the above proteins or
 CC peptide fragments can be used to treat S. pneumoniae infection. The
 CC antibodies can also be used to detect S. pneumoniae cells.
 SQ Sequence 582 BP: 171 A; 98 C; 148 G; 165 T;

Query Match 100.0%; Score 9; DB 50; Length 582;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 279 ttctcagaa 287
 |||||
 Cp 33 ttctcagaa 25

RESULT 125

ID V88046 standard; cDNA; 600 BP.

AC V88046;

DT 12-FEB-1999 (first entry)

DE EST clone FP268.

KW Expressed sequence tag; secreted protein; haematopoiesis regulator;

KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;

KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;

KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.

OS Homo sapiens.

PN M09845437-A2.

PN 15-OCT-1998.

PF 10-APR-1998; U06956.

PF 10-APR-1997; US-837312.

PA (GENE) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,

PI Racie LA, Spaulding V, Treacy M;

DR WPI: 99-070078/06.

PT New polynucleotides encoding human secreted proteins - derived from

PT e.g. human blood, kidney, foetal lung, placenta, testes, brain.

PT ovary, pituitary, retina and colon cDNA libraries

PS Claim 1: Page 264; 641pp; English.

CC The present sequence represents an expressed sequence tag (EST), and is

CC a polynucleotide of the invention. The polynucleotides of the invention

CC are all secreted EST sequences isolated from a variety of human tissue

CC sources. The EST sequences and proteins encoded by them are predicted to

CC have useful biological activities which would make them suitable for

CC treating, preventing or ameliorating medical conditions in humans and

CC animals, although no supporting data is given. Suggested activities

CC include nutritional activity, immune stimulating or suppressing activity,

CC haematopoiesis regulating activity, tissue growth activity,

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition

CC activity. The EST sequences are also stated to be useful for gene

CC therapy.

SQ Sequence 600 BP: 201 A; 157 C; 121 G; 121 T;

Query Match 100.0%; Score 9; DB 54; Length 600;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 ttctcagaa 86
 |||||
 Oy 25 ttctcagaa 33

RESULT 126

ID V21154 standard; cDNA; 603 BP.

AC V21154;

DT 10-JUL-1998 (first entry)

DE 5' nucleotide portion of clone AS186_3.

KW Clone AS186_3; secreted protein; chromosome identification;

KW identification; genetic disorder; nutritional use; cell proliferation;

KW cell differentiation; haematopoiesis regulating activity; tissue growth;

KW anti-inflammatory; tumour inhibition; ds.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 448..603

/*tag= a
 /note= "no stop codon given"

FT M09807855-A2.

PN 26-FEB-1998.

PF 20-AUG-1997; U14641.

PR 23-AUG-1996; US-701931.

PA (GENE) GENETICS INST INC.

PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA,

PI Spaulding V, Treacy M;

DR WPI: 98-169159/15.

DR P-PSDB; W52836.

PT Nucleic acid encoding secreted protein from human cells - useful

PT e.g. as immuno-modulators, anti-tumour agents, promoters of tissue

PT growth, haemostatic and thrombolytic agents etc.

PS Claim 18: Pages 45-46; 63pp; English.

CC The present sequence represents the 5' nucleotide sequence of the clone

CC AS186_3. This clone is a full length clone, isolated from a human fetal

CC kidney cDNA library, using methods which are selective for cDNAs

CC encoding secreted proteins. The products of clone AS186_3 may as

CC tissue/molecular weight markers, for chromosome identification,

CC to identify possible genetic disorders, and to isolate new related DNA.

CC They can also be used for nutritional uses, cytokine and cell

CC proliferation, cell differentiation activity, immune stimulating or

CC suppressing activity, haematopoiesis regulating activity, tissue growth

CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,

CC haemostatic and thrombolytic activity, receptor/ligand activity,

CC anti-inflammatory activity, tumour inhibition activity or other

CC activities.

SQ Sequence 603 BP: 198 A; 95 C; 125 G; 178 T;

Query Match 100.0%; Score 9; DB 42; Length 603;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 416 ttctcagaa 424
 |||||
 Oy 25 ttctcagaa 33

RESULT 127

ID T28030 standard; cDNA; 613 BP.

AC T28030;

DT 31-DEC-1996 (first entry)

DE Mouse incomplete BMP type II receptor kinase (BRK-3) cDNA.

KW BMP type II receptor kinase-3; BRK-3; bone morphogenetic protein;

KW BMP type I receptor kinase; BMP receptor; ds.

OS Mus sp.

FT Key Location/Qualifiers

FT CDS 11..613

FT cds /*tag= a

PN M09614579-A1.

PN 17-MAY-1996.

PF 30-OCT-1995; U14027.

PR 04-NOV-1994; US-334178.

PR 05-JUN-1995; US-462467.

PA (PROC) PROCTER & GAMBLE CO.

PI Rosenbaum JS;

PI WPI: 96-251887/25.

DR P-PSDB; R95234.

PT Assays for bone morphogenetic protein activities - using complex of

PT BMP type I receptor kinase protein and BMP receptor kinase protein

PT BRK-3

PS Disclosure; Page 79-80; 101pp; English.

CC A cDNA clone (T28030) codes for mouse incomplete bone morphogenetic

CC protein (BMP) type II receptor kinase protein-3 (BRK-3) (R95234),

CC corresponding to the extracellular, transmembrane and intracellular

CC juxtaposition region of full-length BRK-3 (R95224). Host cells

CC co-transfected with vectors carrying full-length, incomplete or

CC soluble BRK-3 cDNA and full-length, incomplete or soluble BMP type

CC I receptor kinase protein cDNA (see also T28018-29) express a BMP

CC receptor complex useful for screening cDNAs for BMP receptor

CC affinity or for determining the concentration of a BMP receptor

CC ligand in a clinical sample.

SQ Sequence 613 BP; 160 A; 143 C; 148 G; 162 T;
 Query Match 100.0%; Score 9; DB 24; Length 613;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 88 tctctcagaa 96
 33 tctctcagaa 25
 Cp 33 tctctcagaa 25
 RESULT 128
 ID V89994 standard; CDNA: 614 BP.
 AC V89994.
 DT 15-FEB-1999 (first entry)
 DE EST clone CW775.
 KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibitor; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 OS Homo sapiens.
 PN WO9845436-A2.
 PD 15-OCT-1998.
 PF 10-APR-1998: U06955.
 PR 10-APR-1997: US-838821.
 PA (GENY) GENEKICS INST INC.
 PI Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racle LA, Spaulding V, Treacy M;
 DR WPI: 99-070077/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon CDNA libraries.
 PS Claim 1: Page 397: 618pp: English.
 CC The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 SQ Sequence 614 BP; 144 A; 149 C; 127 G; 194 T;
 Query Match 100.0%; Score 9; DB 53; Length 614;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 245 tctctcagaa 253
 33 tctctcagaa 25
 Cp 33 tctctcagaa 25
 RESULT 129
 ID 099370 standard; CDNA: 627 BP.
 AC 099370.
 DT 08-FEB-1996 (first entry)
 DE Rat allograft inflammatory factor 1 (AIF-1) CDNA.
 KW AIF-1; allograft inflammatory factor 1; transplant rejection;
 KW inhibitor; immunogenic; detection; diagnosis; ds.
 OS Rattus sp.
 PH Key Location/Qualifiers
 FT cds 71..514
 FT /*tag= a
 FT /product= AIF-1_antigenic_peptide
 FT 1..70
 FT /*tag= b
 FT 5'utr 512..627
 FT /*tag= c
 FT 3'utr
 PN WO9517506-A1.

PD 29-JUN-1995. U14724.
 PF 21-DEC-1994; US-171385.
 PR 21-DEC-1993; US-171385.
 PA (HARD) HARVARD COLLEGE.
 PI Russell ME, Utans U;
 DR WPI: 95-240668/31.
 DR P-PSDB: R80521.
 PT DNA encoding allograft rejection factors and immunogenic fragments -
 PT useful for identifying transplant rejection and immunogenic fragments -
 PS Claim 8: Fig 11a: 138pp: English.
 CC 099370 encodes the rat allograft inflammatory factor, AIF-1. The
 CC AIF-1 gene is a differentially expressed allograft gene which is
 CC expressed in allograft tissue during transplant rejection.
 CC Identification of the rat AIF-1 product (R80521) or transcript
 CC indicates that allograft rejection is taking place. The rat AIF-1 gene
 CC and product are therefore useful in the diagnosis of transplant
 CC rejection.
 CC The diagnostic methods used allow rejection (vascular inflammation)
 CC to be detected at an early stage and require only a small amount of
 CC biopsy material.
 CC (Sequence given in sequence ID listing in the specification gives
 CC nucleotide 111 as C)
 SQ Sequence 627 BP; 194 A; 134 C; 169 G; 130 T;
 Query Match 100.0%; Score 9; DB 16; Length 627;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 387 tctctcagaa 395
 33 tctctcagaa 25
 Cp 33 tctctcagaa 25
 RESULT 130
 ID T43455 standard; DNA: 637 BP.
 AC T43455.
 DT 08-AUG-1997 (first entry)
 DE ATM gene exon 15.
 KW ATM gene; ataxia-telangiectasia; progressive genetic disorder; antibody;
 KW central nervous system; immune system; chromosomal instability; therapy;
 KW cancer predisposition; radiation sensitivity; cell cycle abnormality;
 KW multi-system disease; autosomal recessive; cerebellar ataxia; cerebellum;
 KW general motor dysfunction; Purkinje cell; oculocutaneous telangiectasia;
 KW blood vessel; bulbar conjunctiva; facial skin; A-T; ss.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT Intron 1..209
 FT /*tag= a
 FT /number= 14
 FT exon 210..435
 FT /*tag= b
 FT /number= 15
 FT 436..637
 FT /*tag= c
 FT /number= 15
 FT Intron
 PN WO9636691-A1.
 PD 21-NOV-1996.
 PF 16-MAY-1996; U07025.
 PR 16-MAY-1995; US-441822.
 PR 08-APR-1996; US-629001.
 PA (KOHN/) KOHN K I.
 PA (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
 PI Shliah Y;
 DR WPI: 97-012070/01.
 PT New isolated ataxia-telangiectasia gene - used to develop prods. for
 PT the study, diagnosis and treatment of ataxia-telangiectasia.
 PS Claim 1: Page 60: 153pp: English.
 CC T43444-T43496 represent exons of the ATM gene of the invention.
 CC Ataxia-telangiectasia (A-T) is caused by mutations, insertions, or
 CC deletions in the coding region of the ATM gene. A-T is a progressive
 CC genetic disorder affecting the central nervous and immune systems. A-T
 CC involves chromosomal instability, cancer predisposition, radiation
 CC sensitivity, and cell cycle abnormalities. A-T is a multi-system disease

CC inherited in an autosomal recessive manner. Cerebellar ataxia that
 CC gradually develops into general motor dysfunction is the first clinical
 CC hallmark of A-T, and results from progressive loss of Purkinje cells in
 CC the cerebellum. Oculocutaneous telangiectasia (dilation of blood vessels)
 CC develops in the conjunctiva and facial skin, and is later
 CC accompanied by graying of the hair and atrophic changes in the skin. The
 CC co-occurrence of cerebellar ataxia and telangiectases in the conjunctivae
 CC and occasionally on the facial skin (the second early hallmark of A-T)
 CC usually establishes the differential diagnosis of A-T from other
 CC cerebellar ataxias. The ATM gene, can be used in methods for detecting
 CC carriers of a defective gene that causes A-T. The gene can also be used
 CC to generate antibodies. The antibodies and methods can be used in the
 CC study, diagnosis and therapy of A-T.
 SO Sequence 637 BP; 231 A; 88 C; 124 G; 194 T;

Query Match 100.0%; Score 9; DB 31; Length 637;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 244 tttctcagaa 252
 Cp 33 tttctcagaa 25

RESULT 131
 ID 070205 standard; DNA: 666 BP.
 AC 070205: (first entry)
 DT 30-MAR-1995
 DE Influenza haemagglutinin HA2 subunit coding sequence.
 KW Antigen; immunogen; vaccine; influenza; fusion protein; immunity;
 KM haemagglutinin; neuraminidase; flu; ds.
 OS H3N2 A/Victoria influenza virus.
 FS Key
 FT cds
 FT 1.666
 FT /*tag- a
 FT /product- haemagglutinin HA2 subunit.
 PN M09417826-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U01149.
 PR 01-FEB-1993; US-013415.
 PR 18-AUG-1993; US-108914.
 PR 05-NOV-1993; US-149150.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Dillon S, Kane J, Scott M, Shatzman A;
 DR WPI: 94-279392/34.
 DR P-PSDB: R60221.
 PT Vaccines against multi strain influenza virus infection - protect
 PT against influenza A and B.
 PS Disclosure; Page 57-58; 151pp; English.
 CC A vaccine comprising an immunogenic fragment of the HA2 subunit of
 CC the Influenza haemagglutinin (HA) protein from type A subtype IV and
 CC type B IV may be used for stimulating protection in animals against
 CC infection with influenza virus. The vaccine confers multi-strain
 CC immunity against strains IV A and IV B. The vaccines may be
 CC recombinantly produced, optionally as fusion proteins.
 SO Sequence 666 BP; 224 A; 118 C; 166 G; 158 T;

Query Match 100.0%; Score 9; DB 12; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 tttctcagaa 216
 Cp 33 tttctcagaa 25

RESULT 132
 ID 070204 standard; DNA: 666 BP.
 AC 070204:
 DT 30-MAR-1995 (first entry)
 DE Influenza haemagglutinin HA2 subunit coding sequence.
 KW Antigen; immunogen; vaccine; influenza; fusion protein; immunity;
 KM haemagglutinin; neuraminidase; flu; ds.

OS H3N2 A/Udorn influenza virus.
 FS Key
 FT cds
 FT 1.666
 FT /*tag- a
 FT /product- haemagglutinin HA2 subunit.
 PN M09417826-A.
 PD 18-AUG-1994.
 PE 01-FEB-1994; U01149.
 PR 01-FEB-1993; US-013415.
 PR 18-AUG-1993; US-108914.
 PR 05-NOV-1993; US-149150.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Dillon S, Kane J, Scott M, Shatzman A;
 DR WPI: 94-279392/34.
 DR P-PSDB: R60220.
 PT Vaccines against multi strain influenza virus infection - protect
 PT against influenza A and B.
 PS Disclosure; Page 54-55; 151pp; English.
 CC A vaccine comprising an immunogenic fragment of the HA2 subunit of
 CC the Influenza haemagglutinin (HA) protein from type A subtype IV and
 CC type B IV may be used for stimulating protection in animals against
 CC infection with influenza virus. The vaccine confers multi-strain
 CC immunity against strains IV A and IV B. The vaccines may be
 CC recombinantly produced, optionally as fusion proteins.
 SO Sequence 666 BP; 223 A; 117 C; 167 G; 159 T;

Query Match 100.0%; Score 9; DB 12; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 tttctcagaa 216
 Cp 33 tttctcagaa 25

RESULT 133
 ID T95699 standard; DNA: 673 BP.
 AC T95699;
 DT 23-APR-1998 (first entry)
 DE Novel human gene, designated WMP3.
 KW Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;
 KM WW domain; cell signalling; growth regulation; cytoskeleton organisation;
 KW targeted drug screening; modulator; WW domain interaction; WMP3; ss.
 OS Homo sapiens.
 PN M09373723-A1.
 PD 09-OCT-1997.
 PD 03-APR-1997.
 PE 03-APR-1997; U05547.
 PR 03-APR-1996; US-630916.
 PA (CYTO-) CYTOGEN CORP.
 PA (VINC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Kay BK, Pirozzi G;
 DR WPI: 97-503234/46.
 DR P-PSDB: M36796.
 PT Identifying cell signalling and growth regulatory polypeptides by
 PT reaction with multivalent recognition complex - polypeptides are
 PT useful in targeted drug selection
 PS Claim 68; Fig 20; 220p; English.
 CC The present DNA sequence encodes a novel protein WMP3. The WMP3 gene was
 CC identified and isolated from human bone marrow and brain cDNA libraries,
 CC using peptides M38103-05. These peptide recognition units are based on
 CC the sequences of WW domain binding domains of the YAP WW domain binding
 CC proteins WBP-1 and WBP-2. The WW domain is a small functional domain
 CC found in a large number of proteins from a variety of species including
 CC humans, nematodes and yeast. Its name is derived from the observation
 CC that two tryptophan residues, one in the amino terminal portion of the
 CC WW domain and one in the carboxyl terminal portion, are conserved. Most
 CC proteins containing WW domains have a function involving cell signalling
 CC and growth regulation or the organisation of the cytoskeleton.
 CC Polypeptides containing a WW domain are identified by treating a
 CC multivalent recognition unit complex that has selective binding affinity
 CC for a WW domain, with many polypeptides and identifying those with
 CC selective affinity for the complex. Proteins containing WW domains are
 CC used for targeted drug screening, i.e. to identify potential modulators

PA (FEDE/) FEDEROFF H. J.
 PA (FISH/) FISHMAN M. C.
 PA (STRI/) STRITTMATTER SM.
 PA (VALE/) VALENZUELA D.
 PA (ZUBE/) ZUBER M. X.
 PI Federoff HJ, Fishman MC, Strittmatter SM, Valenzuela D;
 PI Zuber MX;
 DR WPI: 94-049278/07.
 PT Recombinant mammalian GAP-43 protein - used to monitor and
 PT regulate neuronal growth in animals, pref. humans
 PS Disclosure: Fig 14; 156pp; English.
 CC The nucleotide sequence of the rat GAP-43 gene promoter region is
 CC given in 057506.
 SQ Sequence 700 BP; 230 A; 87 C; 278 G; 105 T;
 Query Match 100.0%; Score 9; DB 9; Length 700;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 179 ttctgagaa 187
 Oy 25 ttctgagaa 33
 RESULT 137
 ID 053372 standard; DNA: 720 BP.
 AC 053372;
 DT 03-JUN-1994 (first entry)
 DE Sequence encoding glutathione peroxidase.
 KW Glutathione peroxidase; Schistosoma mansoni; schistosoma;
 KW diagnosis; vaccine; ss.
 OS Schistosoma mansoni.
 FH Key location/Qualifiers
 FT CDS 16..525
 FT /*tag- a
 FT /product- Glutathione peroxidase.
 FT misc-difference 142..144
 FT /*tag- b
 FT /note- "This stop codon is ignored due to a
 FT suppressor tRNA inserting a selenocysteine
 FT residue into the amino acid sequence.
 FT misc-difference 304..306
 FT /*tag- C
 FT /transl_except- CAA encodes Glutamic acid.
 PN FR2689906-A.
 PD 15-OCT-1993.
 PR 10-APR-1992: FR-004406.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST. PASTEUR.
 PA (INSP) INST PASTEUR LILLE.
 PI Capron A, Pierce R, Williams D;
 DR WPI: 93-38582/49.
 DR P-PSDB: R44988.
 PT Nucleic acid encoding schistosoma glutathione peroxidase - and
 PT derived peptide, useful in diagnostic reagents and vaccines, also
 PT related vectors, transformed cells, etc.
 PS Claim 2: Page 10: 15pp; French.
 CC The glutathione peroxidase is immunogenic and is used as a
 CC diagnostic reagent and in the preparation of vaccines.
 SQ Sequence 720 BP; 233 A; 129 C; 146 G; 212 T;
 Query Match 100.0%; Score 9; DB 9; Length 720;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 675 ttctcagaa 683
 Cp 33 ttctcagaa 25
 RESULT 138
 ID V34780 standard; DNA: 746 BP.

AC V34780: (first entry)
 DT 11-SEP-1998
 DE Human RATH1.1 DNA.
 KW RATH1.1; RGS gene family; regulator of G-protein signaling; human;
 KW regulation; activation; modulation; signal transduction; T cell;
 KW T helper cell; treatment; autoimmune disorders; arthritis; infection;
 KW graft rejection; asthma; allergy; neoplasm; ss.
 OS Homo sapiens.
 FH Key location/Qualifiers
 FT CDS 39..647
 FT /*tag- a
 FT /product- RATH1.1
 PN M09814579-A1.
 PD 09-APR-1998.
 PR 06-OCT-1997: U18259.
 PR 06-JUN-1997: US-870815.
 PR 04-OCT-1996: US-726228.
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Gameno CJ, Levinson DA;
 DR WPI: 98-240085/21.
 DR P-PSDB: M59294.
 PT New activated T helper cell specific gene, RATH - used for
 PT developing products for treating e.g. autoimmune disorders,
 PT arthritis, graft rejection, asthma, allergy, infections or neoplasms
 PS Claim 1: Fig 2A-B; 127pp; English.
 CC This sequence encodes a novel member of the human regulator of G-protein
 CC signaling (RGS) gene family, RATH1.1. The encoded protein participates
 CC in the regulation, control and/or modulation of G-protein mediated
 CC signal transduction, involved in T cell activation, including T-helper
 CC (TH) cell and TH cell subpopulation activation. The protein can be
 CC used to develop products for treating autoimmune disorders, arthritis,
 CC graft rejection, asthma, allergy, infections or neoplasms.
 SQ Sequence 746 BP; 172 A; 217 C; 215 G; 142 T;
 Query Match 100.0%; Score 9; DB 44; Length 746;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 201 ttctcagaa 209
 Cp 33 ttctcagaa 25
 RESULT 139
 ID T41404 standard; cDNA to mRNA: 756 BP.
 AC T41404;
 DT 13-JAN-1997 (first entry)
 DE Human REST protein DNA binding domain cDNA.
 KW REST protein; REST protein; REST protein; REST protein;
 KW neural cell; differentiation; neurodegenerative disease;
 KW gene therapy; brain cancer; ds.
 OS Homo sapiens.
 PN M09629433-A1.
 PD 26-SEP-1996.
 PR 22-MAR-1996: U039440.
 PR 23-MAR-1995: US-409823.
 PA (UNYX) UNIV NEW YORK STATE RES FOUND.
 PI Chong JA, Mandel G;
 DR WPI: 96-443206/44.
 DR P-PSDB: R99364.
 PT New REST-silencing Transcription factor (REST protein) - used
 PT therapeutically to control differentiation and activity of neural
 PT cells, e.g. in cases of brain cancer
 PS Claim 5: Page 40-42: 109pp; English.
 CC A cDNA clone (T41404) codes for the DNA binding domain (R99364) of
 CC the human REST protein (see also R99365), or REST-silencing
 CC transcription factor, which inhibits expression of neural proteins
 CC in non-neural tissues by silencing promoters contg. REST elements
 CC (see also T41407). The cDNA sequence was derived from a full-length
 CC cDNA clone (T41406) coding for REST obtd. from HeLa cells. It can
 CC be used to produce REST DNA binding domain in transformed host
 CC cells, as a probe, and in gene therapy to treat conditions
 CC associated with excessive neural activity; antisense sequences can

```
CC - be used to treat disorders involving de-differentiated neural cells
SQ sequence 756 BP; 259 A; 164 C; 153 G; 180 T;

Query Match          100.0%   Score 9; DB 24; Length 756;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps

Db      459 ttctcagaa 467
       |||||||
Cp      33 ttctcagaa 25

RESULT 140
ID      T14663 standard; DNA; 776 BP.
AC      T14663;
DT      10-OCT-1996 (first entry)
DE      E6/E7 region of Human Papilloma Virus 16 (HPV 16).
KW      Human papilloma virus; HPV; detection; cervical cancer;
        amplification; hybridisation; diagnosis; transformed cell; E6; E7;
OS      ss.
FH      Key
FT      primer_bind          Location/Qualifiers
FT              1..24
FT              /tag= a
FT              /note= "Primer BB113 binding site."
FT      misc_binding
FT              30..55
FT              /tag= b
FT              /note= "Primer H16-58 binding site."
FT      primer_bind
FT              37..57
FT              /tag= c
FT              /note= "Primer BB4 binding site."
FT      primer_bind
FT              454..474
FT              /tag= d
FT              /note= "Primer BB114 binding site."
FT      primer_bind
FT              480..503
FT              /tag= e
FT              /note= "Primer BB111 binding site."
FT      primer_bind
FT              591..621
FT              /tag= f
FT              /note= "Primer BB109 binding site."
FT      primer_bind
FT              658..681
FT              /tag= g
FT              /note= "Primer H16-686 binding site."
FT      primer_bind
FT              660..683
FT              /tag= h
FT              /note= "Primer BB112 binding site."
FT      primer_bind
FT              715..738
FT              /tag= i
FT              /note= "Primer H16-743 binding site."
FT      primer_bind
FT              745..768
FT              /tag= j
FT              /note= "Primer H16-773 binding site."
FT      primer_bind
FT              745..768
FT              /tag= k
FT              /note= "Primer H16-773 binding site."
PN      US5506105-A.
PD      09-APR-1996.
PR      10-DEC-1991; 808456.
PR      10-DEC-1991; US-808456.
PR      22-MAR-1994; US-216233.
PA      (DADE-) DADE INT INC.
PI      Haydock PV;
PT      WPI: 96-200273/20.
PT      Detection of low copy number intracellular markers - by 3SR
PT      amplification of target RNA in fixed cells then hybridisation with
PT      labelled probe, for detecting human papilloma virus in cervical
PT      cells
PS      Example 2; Figure 3; 21pp; English.
CC      An in situ hybridisation assay for detecting an intracellular marker
CC      of low copy number in cells comprises fixing the cells to a support
CC      using paraformaldehyde; treating the cells with a protease to
CC      permeabilise them without altering morphology; adding amplification
CC      reagents; incubating the cells at below fifty degrees celcius to
CC      perform amplification by self-sustained sequence replication; adding
CC      a labelled probe complementary to the region between the primers;
CC      washing cells to remove unhybridised probe and then detecting the
```

```

CC labelled probe. The E6/E7 region of human papilloma virus (HPV) 16
CC is used especially to detect mRNA being transcribed from this
CC region which becomes active in transformed cells. The method can be
CC used for the early diagnosis of cervical cancer. Primers used to
CC amplify fragments of the E6/E7 region are described in
CC T14664-T14674.
CC
SQ Sequence 776 BP; 260 A; 139 C; 169 G; 208 T;

Query Match 100.0%; Score 9; DB 22; Length 776;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 761 ttctcagaa 769
|||||||
Cp 33 ttctcagaa 25

RESULT 141
ID Q29389 standard; DNA; 790 BP.
AC Q29389;
DT 09-MAR-1993 (first entry)
DE DNA encoding HPV 16 E6/E7 proteins obid. by PCR.
KW Virus vector; vaccinia virus; papillomavirus; HPV; human;
amplification; immunotherapeutic; ss.
OS Human papillomavirus 16.
FH Key Location/Qualifiers
FT cds 1..789
FT FT /*tag- a
FT FT /note- "reading frame 1"
FT FT cds 2..790
FT FT /*tag- b
FT FT /note- "reading frame 2"
FT FT misc.feature 485..781
FT FT /*tag- c
FT FT /note- "second reading frame encoding HPV 16 E7"
FT FT cds 3..788
FT FT /*tag- d
FT FT /note- "reading frame 3"
FT FT misc.feature 6..482
FT FT /*tag- e
FT FT /note- "third reading frame encoding HPV 16 E6"

PN W09216636-A.
PD 01-OCT-1992.
PD 10-MAR-1992; G00424.
PR 14-MAR-1991; GB-005383.
PR (IMMU ) IMMUNOLOGY LTD.
PI Bourasnell MEG, Inglis SC, Munro AJ;
PI WPI; 92-349219/42.
DR P-FDSB; R27723-5.
PT Recombinant virus vectors encoding human papillomavirus proteins
PT -for treating and vaccinating against HPV infections and
PT conditions caused by them, such as cervical cancer
PS Disclousure; Fig 1a; 83bp; English.
PS The fragment of DNA contg. the HPV-16 E6/E7 coding region was
CC prep'd. by PCR from plasmid pBR322/HPV16 (Purst et al., PNAS, 80:
CC 3812 (1983)) using oligonucleotides S05 and S06. The DNA prod. has
CC a site for NcoI at the beginning of the E6 gene and a SmaI site
CC immediately downstream of the termination codon for E7. The E6 and
CC E7 ORFs are fused together to form a single continuous ORF via site
CC directed mutagenesis and the immortalising potential of E7 is removed
CC by altering two key codons of the HPV E7 sequence. The single ORF of
CC HPV-16 E6/E7 may be inserted into vaccinia virus DNA at neutral sites
CC (pref. by inserting two sets of the DNA in opposite orientations to
CC overcome the problem of intertypic recombination) to make a recombinant
CC virus vector for use immunotherapeutically to activate cells of the
CC immune system against HPV. See also Q29385-400 and Q29450-69.
CC Sequence 790 BP; 262 A; 144 C; 175 G; 209 T;

Query Match 100.0%; Score 9; DB 5; Length 790;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 766 ttctcagaa 774

```

CP 33 ttccagaa 25

RESULT 142
ID V17986 standard; DNA: 792 BP.

DE Escherichia coli modified beta-lactamase DNA sequence SEQ ID NO:4.
KW Escherichia coli; Bacillus licheniformis; beta-lactamase; BL gene;
KW modification; identification; sensitive; rapid; cellular assay;
KW drug discovery; modulator; immune response; diabetes; obesity;
KW inflammation; cancer; trauma; ss.

OS Synthetic.
FH Escherichia coli.
FT Key Location/Qualifiers
FT CDS 1..792
FT /tag- a
FT /product- "beta-lactamase"
FT /transl_except- (pos:418..419,aa:Arg)
FT /note- "no stop codon given"

PN WO9813353-A1.
PD 02-APR-1998.
PF 26-SEP-1997; U17395.
PR 26-SEP-1996; US-719697.
PA (AURO-) AURORA BIOSCIENCES CORP.
PI Craig F. Foulkes GJ, Mere L, Negulescu PA, Whitney MA;
DR WPI; 98-230619/20.
DR P-PSDB; W48616.
PT Identifying modulators of genomic nucleic acid - using cell having
PT integrated beta-lactamase construct exposed to test compound then
PT assessed for enzyme expression

PS Example 1; Page 83-84; 112pp; English.
CC The present sequence encodes a modified beta-lactamase (BL) used to
CC exemplify the present invention. The present invention describes a
CC method for identifying proteins or chemicals that directly or indirectly
CC modulate a genomic nucleic acid. The method comprises treating a living
CC cell, which contains a BL-expressing construct integrated into a
CC non-yeast eukaryotic genome, with a predetermined concentration of test
CC compound, then detecting BL activity in the cell. The method is used to
CC identify chemicals such as drugs which are potentially used to treat
CC immune responses; cardiac, vascular, neural, endocrine or
CC gastrointestinal disorders; diabetes; obesity; inflammation; cancer and
CC trauma. Generally the method is used to identify useful/functional
CC regions of the genome; modulators of these regions and cellular
CC pathways. Measurement of BL provides rapid, in vivo, identification/
CC isolation of proteins or chemicals associated with particular biological
CC processes, and characterisation of these can be done in the same cells
CC (even the same assay), eliminating the need for transfer to a secondary
CC screening system.

SO Sequence 792 BP; 216 A; 183 C; 211 G; 182 T;

Query Match 100.0%; Score 9; DB 43; Length 792;
Best Local Similarity 100.0%; Pred. NO. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 ttccagaa 227
| | | | | | | | | |
CP 33 ttccagaa 25

RESULT 143
ID V17983 standard; DNA: 795 BP.
AC V17983:
DT 13-AUG-1998 (first entry)
DE Escherichia coli modified beta-lactamase DNA sequence SEQ ID NO:1.
KW Escherichia coli; Bacillus licheniformis; beta-lactamase; BL gene;
KW modification; identification; sensitive; rapid; cellular assay;
KW drug discovery; modulator; immune response; diabetes; obesity;
KW inflammation; cancer; trauma; ss.
OS Synthetic.
FH Escherichia coli.
FT Key Location/Qualifiers

FT CDS 1..795
FT /tag- a
FT /product- "beta-lactamase"
FT /note- "no stop codon given"

PN WO9813353-A1.
PD 02-APR-1998.
PF 26-SEP-1997; U17395.
PR 26-SEP-1996; US-719697.
PA (AURO-) AURORA BIOSCIENCES CORP.
PI Craig F. Foulkes GJ, Mere L, Negulescu PA, Whitney MA;
DR WPI; 98-230619/20.
DR P-PSDB; W48613.
PT Identifying modulators of genomic nucleic acid - using cell having
PT integrated beta-lactamase construct exposed to test compound then
PT assessed for enzyme expression

PS Example 1; Page 79-80; 112pp; English.
CC The present sequence encodes a modified beta-lactamase (BL) used to
CC exemplify the present invention. The present invention describes a
CC method for identifying proteins or chemicals that directly or indirectly
CC modulate a genomic nucleic acid. The method comprises treating a living
CC cell, which contains a BL-expressing construct integrated into a
CC non-yeast eukaryotic genome, with a predetermined concentration of test
CC compound, then detecting BL activity in the cell. The method is used to
CC identify chemicals such as drugs which are potentially used to treat
CC immune responses; cardiac, vascular, neural, endocrine or
CC gastrointestinal disorders; diabetes; obesity; inflammation; cancer and
CC trauma. Generally the method is used to identify useful/functional
CC regions of the genome; modulators of these regions and cellular
CC pathways. Measurement of BL provides rapid, in vivo, identification/
CC isolation of proteins or chemicals associated with particular biological
CC processes, and characterisation of these can be done in the same cells
CC (even the same assay), eliminating the need for transfer to a secondary
CC screening system.

SO Sequence 795 BP; 215 A; 185 C; 213 G; 182 T;

Query Match 100.0%; Score 9; DB 43; Length 795;
Best Local Similarity 100.0%; Pred. NO. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 222 ttccagaa 230
| | | | | | | | | |
CP 33 ttccagaa 25

RESULT 144
ID V37407 standard; DNA: 813 BP.
AC V37407:
DT 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae coding region.
KW coding region; ORF; open reading frame; antibacterial;
KW infection; prevention; meningitis; ss.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT CDS 197..286
FT /tag- a
FT /product- unknown
PN WO9819689-A1.
PD 14-MAY-1998.
PF 27-OCT-1997; U19226.
PR 01-NOV-1996; US-029930.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PJ;
DR WPI; 98-286586/25.
DR P-PSDB; W61019.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PT of meningitis
PS Claim 1; Page 112-113; 130pp; English.
CC The sequence is that of a coding region isolated from
CC S. pneumoniae. Its encoded protein, or agonists of it,
CC may be useful as an antibacterial for treatment or

CC prevention of infection, specifically caused by *S. pneumoniae*
 CC (particularly meningitis) but possibly also *Helicobacter*
 CC pylori (ulcers and gastric cancer). It may be of particular
 CC use before insertion of an in-dwelling device or any other
 CC invasive procedure. The protein, or nucleic acid encoding
 CC it, can also be used in vaccines to induce a cellular
 CC and/or humoral immune response, or to screen for other
 CC antibacterials. The DNA may also contain flanking sequences
 CC that are potential sources of control elements for bacterial
 CC gene expression. Detecting a sequence encoding the protein
 CC can be used diagnostically, e.g. to detect a mutation for
 CC serotyping or classifying infectious agents.
 SQ Sequence 813 BP: 262 A; 167 C; 163 G; 221 T;

Query Match 100.0%; Score 9; DB 46; Length 813;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 293 tctcagaa 301
 Oy 25 tctcagaa 33

RESULT 145
 ID V40858 standard; DNA: 817 BP.
 AC V40858: 1998 (first entry)
 DE 25-SEP-1998 (first entry)
 DE Beta-lactamase fragment coding sequence.
 KW Beta-lactamase; chimeric enzyme; substrate catalysis modulator; c-erbB2;
 KW binding molecule detection; prostate specific antigen; PSA; CEA;
 KW carcinoma embryonic antigen; ss.
 OS Synthetic.
 FH Key
 FT CDS Location/Qualifiers
 FT 6..815
 FT /*tag= a

PN WO9823731-A2.
 PD 04-JUN-1998.
 PD 26-NOV-1997; IB1643.
 PR 27-NOV-1996; US-757425.
 PA (UYLO-) UNITV CATHOLIQUE LOUVAIN.
 PI Fastrez J, Legendre D, Soumillion P;
 DR WPI: 98-322712/28.
 DR P-PSDB: W57890.

PT New chimeric enzymes, particularly for detection of analytes
 PT comprising a starting enzyme in which a mimotope is inserted to bind
 PT a binding molecule to modulate the activity of the enzyme
 PS Example 1; Fig 1B; 60pp; English.
 CC This sequence encodes a beta-lactamase fragment that can be used in the
 CC chimeric enzyme (CE) of the invention. The CE comprises: (a) a starting
 CC enzyme which is a polypeptide; and (b) a mimotope comprising at least one
 CC amino acid, which mimotope is inserted into the starting enzyme or which
 CC replaces at least one amino acid of the starting enzyme, the insertion or
 CC replacement yielding the CE which has an enzymatic activity of the
 CC starting enzyme and the activity of the CE is modulated upon binding of a
 CC binding molecule (BM) to the mimotope. The BM can bind to the CE to
 CC modulate the catalysis of a substrate by the CE. The CE can be used for
 CC the detection of BMS which can be analyzed such as an antibody specific
 CC for prostate specific antigen (PSA), carcinoma embryonic antigen (CEA),
 CC c-erbB2, products of oncogenes, virus (e.g. HIV or hepatitis), or
 CC bacteria (e.g. *Staphylococcus*). Alternatively the CE can be used for the
 CC detection of ligands of the BMS.
 SQ Sequence 817 BP: 215 A; 190 C; 227 G; 185 T;

Query Match 100.0%; Score 9; DB 45; Length 817;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 221 tctcagaa 229
 Cp 33 tctcagaa 25

RESULT 146

ID V31222 standard; DNA: 827 BP.
 AC V31222:
 DT 01-OCT-1998 (first entry)
 DE E. coli J96 pathogenicity island contig #36.
 KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pher;
 KW PAI V; pher; vaccine; protective immune response; ds.
 OS *Escherichia coli*.
 PN WO982575-A2.
 PD 28-MAR-1998.
 PD 21-NOV-1997; U21347.
 PR 14-OCT-1997; US-061953.
 PR 22-NOV-1996; US-031626.
 PA (HDMN-) HDMN GENOME SCI INC.
 PA (UWMT-) UNITV WISCONSIN.
 PI Chol GH, Dillon PJ, Welch RA;
 DR WPI: 98-314461/27.

PT New isolated uropathogenic E. coli nucleotide sequences - used to
 PT develop products for the detection of pathogenic E. coli and to
 PT elicit an immune response to pathogenic E. coli
 PS Claim 21; Page 137-138; 250pp; English.

CC This sequence represents a E. coli strain J96 contig containing
 CC pathogenicity island (PAI) sequences, and represents a nucleic acid
 CC molecule of the invention. PAIs are large fragments of DNA which comprise
 CC pathogenicity determinants. The sequences of the invention are taken from
 CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pher)
 CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
 CC approximately 94 min (at pher) on the E. coli chromosome and is
 CC approximately 160 kb in size. Antibodies specific to the proteins encoded
 CC by the PAI open reading frames of the invention can be used in kits to
 CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
 CC a protective immune response in an animal to the uropathogenic E. coli
 CC strain J96.
 SQ Sequence 827 BP: 190 A; 161 C; 220 G; 254 T;

Query Match 100.0%; Score 9; DB 46; Length 827;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 tctcagaa 128
 Cp 33 tctcagaa 25

RESULT 147
 ID N80115 standard; DNA: 832 BP.
 AC N80115:
 DT 18-OCT-1990 (first entry)
 DE Human prolactin gene.
 KW Prolactin; milk; contraceptive; dairy cows; lactation; ss.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT signal_peptide 4..87
 FT mat_peptide 88..684
 FT /*tag= a
 FT /*tag= b
 FT /label=prolactin

PN US4725549-A.
 PD 16-FEB-1988.
 PD 23-MAR-1984; 592714.
 PR 22-SEP-1980; US-189160.
 PR 23-MAR-1984; US-592714.
 PA (REGC) University of California.
 PI Cooke NF, Baxter JD;
 DR WPI: 88-070922/10.
 DR P-PSDB: P82079.

PT DNA coding for prolactin - obd. by prepn. of reverse transcript
 PT of mRNA coding for prolactin and inserting into a transfer vector.
 PS Claim 12; Fig 3; 10pp; English.
 CC The cDNA was prepn. from RNA extracted from prolactin secreting
 CC tumours obd. after surgical removal. The sequence can be inserted
 CC into expression vectors for the prodn. of prolactin which can be
 CC admin. to dairy cows to increase milk yield. It can also be used
 CC as a female contraceptive and to ensure adequate milk prodn. for

```
CC breast feeding mothers.
CC See also N80114.
SQ Sequence 832 BP; 216 A; 233 C; 182 G; 201 T;

Query Match 100.0%; Score 9; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 709 tctctagaa 717
    |||||
Oy 25 tctctagaa 33

RESULT 148
ID V84414 standard: DNA: 843 BP.
AC V84414;
DT 01-MAR-1999 (first entry)
DE Human secreted protein gene 4 clone HLRE125.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN W09854963-A2.
PD 10-DEC-1998.
PE 04-JUN-1998; U11422.
PF 18-DEC-1997; US-070923.
PR 06-JUN-1997; US-048877.
PR 06-JUN-1997; US-048881.
PR 06-JUN-1997; US-048884.
PR 06-JUN-1997; US-048893.
PR 06-JUN-1997; US-048896.
PR 06-JUN-1997; US-048899.
PR 06-JUN-1997; US-048915.
PR 06-JUN-1997; US-048949.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048972.
PR 06-JUN-1997; US-049020.
PR 06-JUN-1997; US-049375.
PR 05-SEP-1997; US-057628.
PR 05-SEP-1997; US-057635.
PR 05-SEP-1997; US-057644.
PR 05-SEP-1997; US-057647.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057661.
PR 05-SEP-1997; US-057667.
PR 05-SEP-1997; US-057761.
PR 05-SEP-1997; US-057764.
PR 05-SEP-1997; US-057770.
PR 05-SEP-1997; US-057775.
PR 05-SEP-1997; US-057778.
PR 06-JUN-1997; US-048875.
PR 06-JUN-1997; US-048878.
PR 06-JUN-1997; US-048882.
PR 06-JUN-1997; US-048885.
PR 06-JUN-1997; US-048894.
PR 06-JUN-1997; US-048897.
PR 06-JUN-1997; US-048900.
PR 06-JUN-1997; US-048916.
PR 06-JUN-1997; US-048962.
PR 06-JUN-1997; US-048970.
PR 06-JUN-1997; US-048974.
PR 06-JUN-1997; US-049373.
PR 05-SEP-1997; US-057584.
PR 05-SEP-1997; US-057629.
PR 05-SEP-1997; US-057642.
PR 05-SEP-1997; US-057645.
PR 05-SEP-1997; US-057648.
PR 05-SEP-1997; US-057651.

PR 05-SEP-1997; US-057662.
PR 05-SEP-1997; US-057668.
PR 05-SEP-1997; US-057762.
PR 05-SEP-1997; US-057765.
PR 05-SEP-1997; US-057771.
PR 05-SEP-1997; US-057776.
PR 06-JUN-1997; US-048876.
PR 06-JUN-1997; US-048880.
PR 06-JUN-1997; US-048883.
PR 06-JUN-1997; US-048892.
PR 06-JUN-1997; US-048895.
PR 06-JUN-1997; US-048898.
PR 06-JUN-1997; US-048901.
PR 06-JUN-1997; US-048917.
PR 06-JUN-1997; US-048963.
PR 06-JUN-1997; US-048971.
PR 06-JUN-1997; US-049019.
PR 06-JUN-1997; US-049374.
PR 05-SEP-1997; US-057627.
PR 05-SEP-1997; US-057634.
PR 05-SEP-1997; US-057643.
PR 05-SEP-1997; US-057646.
PR 05-SEP-1997; US-057649.
PR 05-SEP-1997; US-057654.
PR 05-SEP-1997; US-057666.
PR 05-SEP-1997; US-057760.
PR 05-SEP-1997; US-057763.
PR 05-SEP-1997; US-057769.
PR 05-SEP-1997; US-057774.
PR 05-SEP-1997; US-057777.
PR (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Epner R, Endress GA,
PI Fan P, Feng P, Ferrle AM, Fischer CL, Florence C,
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DM,
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
PI WPI: 99-059665/05.
DR P-PSDB: W88537.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 4; Page 273; 772pp; English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischaemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
SQ Sequence 843 BP; 244 A; 171 C; 152 G; 267 T;

Query Match 100.0%; Score 9; DB 56; Length 843;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Db . 547 ttctcagaa 555
 |||||||||
 CC 33 ttctcagaa 25

RESULT 149
 ID 048592 standard; DNA: 850 BP.
 AC 048592:
 DT 16-FEB-1994 (first entry)
 DE Autonomously replicating sequence.
 KW Expression cassette; expression vector; inducible; methanol;
 OS glycerol; yeast; ss.
 OS Methylophilic yeast.
 PN BP-558024-A.
 PD 01-SEP-1993
 PF 26-FEB-1993; 103040.
 PR 28-FEB-1992; JP-043361.
 PA (SUNR) SUNTORY LTD.
 PI Hatanaka H, Kondo H, Sakai Y, Shibano Y, Tani Y;
 DR WPI: 93-274483/35.
 PT New expression cassettes inducible by methanol and/or glycerol -
 having promoter and terminator derived from alcohol oxidase gene
 PS Claim 6; Page 22; 37pp; English.
 CC An expression cassette which contains the alcohol oxidase promoter
 CC is induced when in the presence of methanol/glycerol. Under such
 CC conditions any heterologous gene located downstream of the promoter
 CC will be expressed. The expression cassette can be used to construct
 CC an expression vector which in turn can be used to transform cells.
 CC When transformed cells are cultured in the presence of methanol/
 CC glycerol expression of the heterologous gene is induced such that
 CC large quantities of the desired product is produced. The
 CC transformed cells are used particularly for the production of
 CC enzymes such as adenylate kinases, cytochrome C and peroxidases.
 CC The autonomously replicating sequence is used in the construction
 CC of the expression vector, conferring the ability of replication
 CC on the expression vector.
 SQ Sequence 850 BP; 351 A; 83 C; 121 G; 295 T;

Query Match 100.0%; Score 9; DB 8; Length 850;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 337 ttctcagaa 345
 |||||||||
 QY 25 ttctcagaa 33

RESULT 150
 ID T39913 standard; DNA: 852 BP.
 AC T39913:
 DT 23-JAN-1997 (first entry)
 DE Maize acetyl COA carboxylase Type A2 genomic clone F.
 KW Acetyl COA carboxylase; ACCase; herbicide tolerance;
 KW cyclohexanedione; arylxyphenoxypyrolic acid; vegetable oil;
 KW oilseed; maize; corn; ss.
 OS Zea mays line B73.
 PN W09631609-A2.
 PD 10-OCT-1996
 PF 04-APR-1996; U04625.
 PR 05-APR-1995; US-417089.
 PA (MINU) UNIV MINNESOTA.
 PI Egli MA, Gengenbach BG, Gronwald JW, Lutz SM, Somers DA;
 DR WPI: 96-465030/46.
 PT DNA encoding maize acetyl coenzyme A carboxylase gene - used for
 PT prodn. of plants with herbicide tolerance or altered oil content
 PS Claim 2; Page 92; 13pp; English.
 CC 6 Genomic DNA fragments (T39908-13) of maize Type A2 acetyl COA
 CC carboxylase (ACCase) were isolated from a B37 genomic library by
 CC screening with a 2 kb subclone of cDNA clone #15-14 (see also
 CC T39904). Type A1 (see also T39906-07) and Type B (T39914-16)
 CC genomic clones and a full-length cDNA clone (T39905) for ACCase
 CC (W05590) have also been obtd. ACCase sequences can be incorporated

CC into vectors and used to increase the herbicide tolerance or oil
 CC content of transgenic plants, or to produce ACCase in heterologous
 CC systems.
 CC Sequence 852 BP; 238 A; 165 C; 193 G; 254 T;

Query Match 100.0%; Score 9; DB 24; Length 852;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 560 ttctcagaa 568
 |||||||||
 QY 25 ttctcagaa 33

RESULT 151
 ID V17984 standard; DNA: 858 BP.
 AC V17984:
 DT 13-AUG-1998 (first entry)
 DE Escherichia coli modified beta-lactamase DNA sequence SEQ ID NO:2.
 KW Escherichia coli; Bacillus licheniformis; beta-lactamase; BL gene;
 KW modification; identification; sensitive; rapid; cellular assay;
 KW drug discovery; modulator; immune response; diabetes; obesity;
 KW inflammation; cancer; trauma; ss.
 OS Synthetic.
 OS Escherichia coli.
 FH Key
 FT Location/Qualifiers
 FT 1..858
 FT CDS
 FT /*tag= a
 FT /product= "beta-lactamase"
 FT /note= "no stop codon given"
 PN W09813353-A1.
 PD 02-APR-1998.
 PF 26-SEP-1997; U17395.
 PR 26-SEP-1996; US-719697.
 PA (AURO) AURORA BIOSCIENCES CORP.
 PI Craig F, Foulkes GJ, Mere L, Negulescu PA, Whitney MA;
 DR WPI: 98-230619/20.
 DR P-PSDB: W48614.

PT Integrated beta-lactamase construct exposed to test compound then
 PT assayed for enzyme expression
 PS Example 1; Page 80-81; 11pp; English.
 CC The present sequence encodes a modified beta-lactamase (BL) used to
 CC exemplify the present invention. The present invention describes a
 CC method for identifying proteins or chemicals that directly or indirectly
 CC modulate a genomic nucleic acid. The method comprises treating a living
 CC cell, which contains a BL-expressing construct integrated into a
 CC non-yeast eukaryotic genome, with a predetermined concentration of test
 CC compound, then detecting BL activity in the cell. The method is used to
 CC identify chemicals such as drugs which are potentially used to treat
 CC immune responses; cardiac, vascular, neural, endocrine or
 CC gastrointestinal disorders; diabetes; obesity; inflammation; cancer and
 CC trauma. Generally the method is used to identify useful/functional
 CC regions of the genome; modulators of these regions and cellular
 CC pathways. Measurement of BL provides rapid, in vivo, identification/
 CC isolation of proteins or chemicals associated with particular biological
 CC processes, and characterisation of these can be done in the same cells
 CC (even the same assay), eliminating the need for transfer to a secondary
 CC screening system.
 SQ Sequence 858 BP; 222 A; 202 C; 223 G; 211 T;

Query Match 100.0%; Score 9; DB 43; Length 858;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttctcagaa 293
 |||||||||
 Cp 33 ttctcagaa 25

RESULT 152
 ID V59591 standard; DNA: 862 BP.
 AC V59591:

06-JAN-1999 (first entry)
 DE Human secreted protein gene 81 clone HMMN32.
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; cancer; neurodegenerative disorder; leukemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 PN WO9839448-A2.
 PD 11-SEP-1998.
 PE 06-MAR-1998: U04493.
 PR 02-OCT-1997: US-061060.
 PR 07-MAR-1997: US-038621.
 PR 07-MAR-1997: US-040161.
 PR 07-MAR-1997: US-040162.
 PR 07-MAR-1997: US-040163.
 PR 07-MAR-1997: US-040333.
 PR 07-MAR-1997: US-040334.
 PR 07-MAR-1997: US-040336.
 PR 07-MAR-1997: US-040626.
 PR 11-APR-1997: US-043311.
 PR 11-APR-1997: US-043312.
 PR 11-APR-1997: US-043313.
 PR 11-APR-1997: US-043314.
 PR 11-APR-1997: US-043568.
 PR 11-APR-1997: US-043569.
 PR 11-APR-1997: US-043576.
 PR 11-APR-1997: US-043578.
 PR 11-APR-1997: US-043580.
 PR 11-APR-1997: US-043669.
 PR 11-APR-1997: US-043670.
 PR 11-APR-1997: US-043671.
 PR 11-APR-1997: US-043672.
 PR 11-APR-1997: US-043674.
 PR 23-MAY-1997: US-047492.
 PR 23-MAY-1997: US-047500.
 PR 23-MAY-1997: US-047501.
 PR 23-MAY-1997: US-047502.
 PR 23-MAY-1997: US-047503.
 PR 23-MAY-1997: US-047581.
 PR 23-MAY-1997: US-047582.
 PR 23-MAY-1997: US-047583.
 PR 23-MAY-1997: US-047584.
 PR 23-MAY-1997: US-047585.
 PR 23-MAY-1997: US-047586.
 PR 23-MAY-1997: US-047587.
 PR 23-MAY-1997: US-047588.
 PR 23-MAY-1997: US-047589.
 PR 23-MAY-1997: US-047590.
 PR 23-MAY-1997: US-047592.
 PR 23-MAY-1997: US-047593.
 PR 23-MAY-1997: US-047594.
 PR 23-MAY-1997: US-047595.
 PR 23-MAY-1997: US-047596.
 PR 23-MAY-1997: US-047597.
 PR 23-MAY-1997: US-047598.
 PR 23-MAY-1997: US-047599.
 PR 23-MAY-1997: US-047600.
 PR 23-MAY-1997: US-047601.
 PR 23-MAY-1997: US-047612.
 PR 23-MAY-1997: US-047613.
 PR 23-MAY-1997: US-047614.
 PR 23-MAY-1997: US-047615.
 PR 23-MAY-1997: US-047617.
 PR 23-MAY-1997: US-047618.
 PR 23-MAY-1997: US-047632.
 PR 23-MAY-1997: US-047633.
 PR 06-JUN-1997: US-048964.
 PR 06-JUN-1997: US-048974.
 PR 13-JUN-1997: US-049610.

PR 08-JUL-1997: US-051926.
 PR 16-JUL-1997: US-052874.
 PR 18-AUG-1997: US-055724.
 PR 22-AUG-1997: US-056630.
 PR 22-AUG-1997: US-056631.
 PR 22-AUG-1997: US-056632.
 PR 22-AUG-1997: US-056636.
 PR 22-AUG-1997: US-056637.
 PR 22-AUG-1997: US-056662.
 PR 22-AUG-1997: US-056664.
 PR 22-AUG-1997: US-056845.
 PR 22-AUG-1997: US-056862.
 PR 22-AUG-1997: US-056864.
 PR 22-AUG-1997: US-056872.
 PR 22-AUG-1997: US-056874.
 PR 22-AUG-1997: US-056875.
 PR 22-AUG-1997: US-056876.
 PR 22-AUG-1997: US-056877.
 PR 22-AUG-1997: US-056878.
 PR 22-AUG-1997: US-056879.
 PR 22-AUG-1997: US-056880.
 PR 22-AUG-1997: US-056881.
 PR 22-AUG-1997: US-056882.
 PR 22-AUG-1997: US-056884.
 PR 22-AUG-1997: US-056886.
 PR 22-AUG-1997: US-056887.
 PR 22-AUG-1997: US-056888.
 PR 22-AUG-1997: US-056889.
 PR 22-AUG-1997: US-056892.
 PR 22-AUG-1997: US-056893.
 PR 22-AUG-1997: US-056894.
 PR 22-AUG-1997: US-056903.
 PR 22-AUG-1997: US-056908.
 PR 22-AUG-1997: US-056909.
 PR 22-AUG-1997: US-056910.
 PR 22-AUG-1997: US-056911.
 PR 05-SEP-1997: US-057630.
 PR 05-SEP-1997: US-057669.
 PR 05-SEP-1997: US-057761.
 PR 12-SEP-1997: US-058785.
 PA (HDMN-) HDMN GENE SCI INC.
 PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,
 PI Feng P, Ferris AM, Fischer CL, Florence KA, Greene JM, Hu JS,
 PI Kyaw H, Latleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 DR WPI: 98-506364/43.
 DR P-PSDB: W74811.
 PT New isolated human genes and the secreted polypeptide(s) they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 PS Claim 1: Page 312; 721pp; English.
 CC This sequence represents a nucleic acid molecule designated Gene 81 from
 CC the human cDNA clone HMMN32 (deposited as clone ATCC 97900 and ATCC
 CC 209046) which encodes a secreted human protein. The gene can be used to
 CC generate fusion proteins by linking to the gene to a human immunoglobulin
 CC Fc portion (e.g. V59502) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 186 polynucleotides, based on
 CC which tissues they are most highly expressed in (see V59511 for described
 CC uses).
 SQ Sequence 862 BP; 193 A; 203 C; 211 G; 254 T;
 Query Match 100.0%; Score 9; DB 51; Length 862;
 Best Local Similarity 100.0%; Pred. No. 1,37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 371 ttctcagaa 379

CP 33 ttctcgagaa 25

RESULT 153
ID V29323 standard; DNA: 867 BP.

AC V29323;

DT 30-JUL-1998 (first entry)

DE Type A2 Accase genomic clone D partial nucleotide sequence.

KW ACCase; maize; herbicide resistant; corn plant; tolerance; Acc1; Acc2;

KW cyclohexanedione; CHD; arylloxypheoxypropionic acid; APA; enzyme; ss.

OS zea mays.

PN MO9808963-A1.

PD 05-MAR-1998.

PF 29-AUG-1997; U15344.

PR 30-AUG-1996; US-697826.

PI Egil MA, Gengenbach BG, Lutz SM, Marshall LC, Parker WB,

PI Somers DA, Vandee KL, Wyse DL;

DR WPI: 98-207043/18.

PT Herbicide resistant corn plants - prepared using Acc1 and Acc2 gene

PT combinations

PS Example 7; Fig 7D; 112pp; English.

CC This is the Type A2 Accase genomic clone D partial nucleotide sequence.

CC The invention provides methods of preparing an herbicide resistant corn

CC plant which comprises crossing a first corn plant to a second corn plant

CC so as to yield a progeny plant, where the first corn plant comprises at

CC least 1 herbicide resistant allele and the second plant comprises at

CC least 1 herbicide resistant allele which is not allelic to the herbicide

CC resistant allele present in the first plant. The herbicide resistant corn

CC plants are prepared using Acc1 and Acc2 gene combinations which impart

CC cyclohexanedione (CHD) or arylloxypheoxypropionic acid (APA) herbicide

CC tolerance to the corn plant. The methods are used to impart CHD and APA

CC herbicide tolerance to corn plants and to produce CHD or APA herbicide

CC tolerant zea mays (corn) homozygous or heterozygous for Acc1 and

CC homozygous or heterozygous for Acc2. The methods can also be used to

CC impart tolerance to a corn plant to an agent which inhibits acetyl CoA

CC carboxylase, selected from 3-(2,4-dichlorophenyl)-perhydropyridazine-2,

CC 4-one, 3-isopropyl-6-(N-12,2-dimethylpropyl)-acetamido-1,3,5-triazine

CC -2,4 (1H,3H)dione, sorphen A and their structural analogues.

SO Sequence: 867 BP; 240 A; 171 C; 198 G; 256 T;

Query Match 100.0%; Score 9; DB 42; Length 867;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 576 ttctcgagaa 584

OY 25 ttctcgagaa 33

RESULT 154

ID T72682 standard; DNA: 878 BP.

AC T72682;

DT 16-SEP-1997 (first entry)

DE PIK-related kinase MCCSI riboprobe.

KW Mammalian cell cycle surveillance 1; MCCSI; PIK-related kinase;

KW phosphatidylinositol kinase; DNA repair; ataxia telangiectasia;

KW riboprobe; ss.

OS Synthetic.

PN MO9718323-A2.

PD 22-MAY-1997.

PF 18-NOV-1996; U19337.

PR 16-NOV-1995; US-558666.

PR 27-FEB-1996; US-007312.

PR 21-OCT-1996; US-725304.

PA (ICOS-) ICOS CORP.

PI Hoekstra MF, Holtzman DA, Keegan KS;

DR WPI: 97-289299/26.

PT Phosphatidylinositol kinase-related mammalian cell cycle

PT surveillance 1 alpha and beta proteins - useful for identifying

PT modulators and inhibitors for treating, e.g. ataxia telangiectasia

PS Example 8; Page 130; 212pp; English.

CC A DNA sequence (T72682) corresponds to a riboprobe generated from

CC murine mammalian cell cycle surveillance 1 (MCCSI) DNA by in vitro

CC RNA transcription. It was used to examine expression of MCCSI

CC (see also W19723-24) mRNA and protein in normal and irradiated

CC mouse testes and in mouse embryos by in situ hybridisation. In

CC both normal and irradiated mouse testes signal was observed in the

CC cytoplasm of spermatogonia and spermatocytes. The expression level

CC in irradiated testis was not increased over that seen in normal

CC testis.

SO Sequence 878 BP; 284 A; 160 C; 195 G; 239 T;

Query Match 100.0%; Score 9; DB 32; Length 878;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 228 ttctcgagaa 236

OY 25 ttctcgagaa 33

RESULT 155

ID Q43705 standard; DNA: 890 BP.

AC Q43705;

DT 30-SEP-1993 (first entry)

DE Sequence of the beta-lactamase gene from Escherichia coli plasmid

DE pBR322.

KW Beta-lactamase; gene; ss.

OS Escherichia coli.

FT Key

FT cds

FT Location/Qualifiers

FT 19..879

FT /*tag= a

PN MO9311242-A.

PD 10-JUN-1993.

PR 29-NOV-1992; BE0052.

PR 29-NOV-1991; BE-001106.

PA (UPLD-) UNIV CATHOLIQUE LOUVAIN.

PI Faetzel J;

DR WPI: 93-197060/24.

DR P-PSDB: R37994.

PT Selecting recombinant microorganisms expressing enzymes on the

PT surface - by affinity chromatography, using specific active site

PT inhibitor; partic. for isolating phage(s)

PS Example; Fig 1; 45pp; French.

CC The nucleotide sequence of the gene for beta-lactamase on E. coli

CC plasmid pBR322 which is reproduced in Figure 1 is published in

CC Suckliffe, PNAS USA, 75, p. 3737 to 3741 (1987).

SO Sequence 890 BP; 238 A; 204 C; 229 G; 219 T;

Query Match 100.0%; Score 9; DB 7; Length 890;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 303 ttctcgagaa 311

CP 33 ttctcgagaa 25

RESULT 156

ID V32990 standard; cDNA: 892 BP.

AC V32990;

DT 20-NOV-1998 (first entry)

DE Human p24 vesicle trafficking protein Hp24-2 nucleotide sequence.

KW ss; human; p24 vesicle-trafficking protein; endocrinal; cancer;

KW secretory disease; inflammatory disorder; gastrointestinal disorder.

OS Homo sapiens.

FT Key

FT CDS

FT Location/Qualifiers

FT 88..777

FT /*tag= a

FT /product= "hp24-2"

PN MO9836068-A2.

PD 20-AUG-1998.

PR 12-FEB-1998; U02813.

PR 14-FEB-1997; US-801740.

PA (INCY-) INCYTE PHARM INC.
 PI Bandman O. GOLI SK:
 DR WPI: 98-467174/40.
 DR P-PSDB: W49100.
 PT New human p24 vesicle-trafficling protein and related nucleic acid -
 PT useful for diagnosis, treatment and prevention of e.g. cancer,
 PT cystic fibrosis, diabetes etc.
 PS Claim 22: Fig 2: 69pp: English.
 CC The human p24 vesicle-trafficling proteins and agonists, optionally
 CC expressed in vivo from, are used to treat or prevent diseases associated
 CC with abnormal vesicle trafficking, including endocytosis, secretory,
 CC inflammatory or gastrointestinal disorders. Antagonists of human
 CC p24, e.g. antisense nucleic acid, ribozymes, triplex-forming molecules,
 CC antibodies, are used to treat or prevent cancers. Antibodies may also be
 CC used to diagnose the specified conditions (as immunoassay reagents), to
 CC monitor treatment, in competitive drug screens and to purify the
 CC protein from natural sources.
 SQ Sequence 892 BP: 268 A: 192 C: 210 G: 221 T:

Query Match 100.0% Score 9: DB 48: Length 892:
 Best Local Similarity 100.0%: Pred. No. 1.37e+03:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 432 ttctcgagaa 440
 |||||
 OY 25 ttctcgagaa 33

RESULT 157
 ID V06143 standard: DNA: 911 BP.
 AC V06143:
 DT 05-MAY-1998 (first entry)
 DE Viral infection: tumour suppressor: cellular gene: rat: cancer:
 KW serum protein: inhibitor: malignant phenotype: HIV: influenza:
 KM hepatitis: retrovirus: immunodeficiency: ds.
 OS Rattus sp.
 PN WO9739119-A1.
 PD 23-OCR-1997.
 PR 11-APR-1997: U06067.
 PR 15-APR-1996: US-015334.
 PA (UYVA-) UNIV VANDERBILT.
 PI Dubois RN, Organ EL, Rubin DH;
 DR WPI: 97-526456/48.
 PT Genes involved in viral infection and tumour suppression - used to
 PT develop products for reducing or preventing viral infection or for
 PT suppressing tumours.
 PS Claim 1: Page 75: 101pp: English.
 CC The present sequence represents a viral infection gene. The present
 CC invention describes nucleic acid sequences isolated from rat. The
 CC sequences of the invention comprise 70 viral infection (VI) genes and
 CC 8 tumour suppressor (TS) genes. Propagating cell cultures in the
 CC absence of the serum protein (SP) allows selective elimination of cells
 CC persistently infected with a virus from the cell culture. Inhibitors of
 CC the SP can be used for reducing or inhibiting a viral infection.
 CC Inhibitors of the TS gene products can be used to suppress a malignant
 CC phenotype (MP). The methods and inhibitors can be used with viruses
 CC such as HIV, influenza, hepatitis virus or animal retroviruses such as
 CC simian immunodeficiency virus, avian immunodeficiency virus, bovine
 CC immunodeficiency virus, feline immunodeficiency virus, equine infectious
 CC anemia virus, caprine arthritis encephalitis virus or vesna virus.
 CC Because the identified genes are non-essential to cell survival, the
 CC treatment methods can be used in subjects without serious detrimental
 CC effects to the subjects.
 SQ Sequence 911 BP: 274 A: 131 C: 199 G: 279 T:

Query Match 100.0% Score 9: DB 38: Length 911:
 Best Local Similarity 100.0%: Pred. No. 1.37e+03:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 495 ttctcgagaa 503
 |||||
 CP 33 ttctcgagaa 25

RESULT 158
 ID Q70208 standard: DNA: 918 BP.
 AC Q70208;
 DT 30-MAR-1995 (first entry)
 DE Sequence encoding immunogenic fragment of influenza haemagglutinin.
 KW Antigen: immunogen: vaccine: influenza: fusion protein: immunity;
 KM haemagglutinin: neuraminidase; flu: ds.
 OS Influenza virus.
 FH Key
 FT Location/Qualifiers
 FT cds
 FT 1..918
 FT /tag- a
 FT /product- NS1(1-81)H3HA2(1-221).
 FT /note- "The sequence encodes a fusion protein of
 FT which the N-terminal 81 amino acids are
 FT derived from the influenza NS1 protein and
 FT the remainder of the sequence is comprised
 FT of amino acids 1-221 of the HA2 subunit of
 FT the H3HA2 subtype of influenza."
 PN WO9417826-A.
 PD 18-AUG-1994.
 PR 01-FEB-1994: U01149.
 PR 01-FEB-1993: US-013415.
 PR 18-AUG-1993: US-108914.
 PR 05-NOV-1993: US-149150.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Dillon S, Kane J, Scott M, Shatzman A;
 DR WPI: 94-279392/34.
 DR P-PSDB: R60194.
 PT Vaccines against multi strain influenza virus infection - protect
 PT against influenza A and B
 PS Example 10: Page 112: 151pp: English.
 CC A vaccine comprising an immunogenic fragment of the HA2 subunit of
 CC the influenza haemagglutinin (HA) protein from type A subtype IV and
 CC type B IV may be used for stimulating protection in animals against
 CC infection with influenza virus. The vaccine confers multi-strain
 CC immunity against strains IV A and IV B. The vaccines may be
 CC recombinantly produced, optionally as fusion proteins. NOTE: This
 CC sequence is a modified version of the sequence given in Q70190,
 CC having cytosine at positions 622, 625 and 634 instead of adenine, and
 CC having thymine at positions 624, 627 and 636 instead of guanine.
 CC These alterations are silent mutations 1.e. the encoded polypeptide
 CC is unaltered. The modifications were made to increase the yield of
 CC the encoded H3HA2 protein.
 SQ Sequence 918 BP: 292 A: 179 C: 231 G: 216 T:

Query Match 100.0% Score 9: DB 12: Length 918:
 Best Local Similarity 100.0%: Pred. No. 1.37e+03:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 463 ttctcgagaa 471
 |||||
 CP 33 ttctcgagaa 25

RESULT 159
 ID Q70190 standard: DNA: 918 BP.
 AC Q70190;
 DT 28-MAR-1995 (first entry)
 DE Sequence encoding immunogenic fragment of influenza haemagglutinin.
 KW Antigen: immunogen: vaccine: influenza: fusion protein: immunity;
 KM haemagglutinin: neuraminidase; flu: ds.
 OS Influenza virus.
 FH Key
 FT Location/Qualifiers
 FT cds
 FT 1..918
 FT /tag- a
 FT /product- NS1(1-81)H3HA2(1-221).
 FT /note- "The sequence encodes a fusion protein of
 FT which the N-terminal 81 amino acids are
 FT derived from the influenza NS1 protein and
 FT the remainder of the sequence is comprised
 FT of amino acids 1-221 of the HA2 subunit of

the H3HA2 subtype of Influenza."

FT W09417826-A.
 PD 18-AUG-1994.
 PR 01-FEB-1994; U01149.
 PR 01-FEB-1993; US-013415.
 PR 18-AUG-1993; US-108914.
 PR 05-NOV-1993; US-149150.
 PA (SMK) SMITHKLINE BECKHAM CORP.
 PI Dillon S, Kane J, Scott M, Shatzman A;
 DR WPI: 94-279392/34.
 DR P-PDB: R60194.
 PT Vaccines against multi strain Influenza virus infection - protect
 PT against Influenza A and B
 PS Claim 23; Page 63-64; 15pp; English.
 CC A vaccine comprising an immunogenic fragment of the HA2 subunit of
 CC the Influenza haemagglutinin (HA) protein from type A subtype IV and
 CC type B IV may be used for stimulating protection in animals against
 CC infection with Influenza virus. The vaccine confers multi-strain
 CC immunity against strains IV A and IV B. The vaccines may be
 CC recombinantly produced, optionally as fusion proteins.
 SQ Sequence 918 BP; 295 A; 176 C; 234 G; 213 T;

Query Match 100.0%; Score 9; DB 12; Length 918;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 463 ttctcagaa 471
 CP 33 ttctcagaa 25

RESULT 160

ID V06123 standard; DNA; 924 BP.
 AC V06123.
 DT 05-MAY-1998 (first entry)
 DE Viral infection gene SEQ ID NO:43.
 KW Viral infection; tumour suppressor; cellular gene; rat; cancer;
 KW serum protein; inhibitor; malignant phenotype; HIV; Influenza;
 KW hepatitis; retrovirus; immunodeficiency; ds.
 OS Rattus sp.
 PN W09739119-A1.
 PD 23-OCT-1997.
 PE 11-APR-1997; U06067.
 PR 15-APR-1996; US-015334.
 PA (UYVA-) UNIV VANDERBILT.
 PI Dubois RN, Organ EL, Rubin DH;
 DR WPI: 97-528456/48.
 PT Genes involved in viral infection and tumour suppression - used to
 PT develop products for reducing or preventing viral infection or for
 PT suppressing tumours
 PS Claim 1; Page 62; 101pp; English.
 CC The present sequence represents a viral infection gene. The present
 CC invention describes nucleic acid sequences isolated from rat. The
 CC sequences of the invention comprise 70 viral infection (VI) genes and
 CC 8 tumour suppressor (TS) genes. Propagating cell cultures in the
 CC absence of the serum protein (SP) allows selective elimination of cells
 CC persistently infected with a virus from the cell culture. Inhibitors of
 CC the SP can be used for reducing or inhibiting a viral infection.
 CC Inhibitors of the TS gene products can be used to suppress a malignant
 CC phenotype (MP). The methods and inhibitors can be used with viruses
 CC such as HIV, Influenza, hepatitis virus or animal retroviruses such as
 CC simian immunodeficiency virus, avian immunodeficiency virus, bovine
 CC immunodeficiency virus, feline immunodeficiency virus, equine infectious
 CC anemia virus, caprine arthritis encephalitis virus or visna virus.
 CC Because the identified genes are non-essential to cell survival, the
 CC treatment methods can be used in subjects without serious detrimental
 CC effects to the subjects.
 SQ Sequence 924 BP; 274 A; 133 C; 202 G; 284 T;

Query Match 100.0%; Score 9; DB 38; Length 924;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 508 ttctcagaa 516
 CP 33 ttctcagaa 25

RESULT 161

ID N91055 standard; DNA; 953 BP.
 AC N91055.
 DT 05-JUL-1990 (first entry)
 DE Sequence encoding exon K and portions of flanking introns of novel
 DE amyloid precursor protein (NAP).
 KW SPAP; amyloid precursor protein; dysbolism; INS76; ds.
 OS Homo sapiens.
 FH Key
 FT exon.
 FT Location/Qualifiers
 FT 623..756
 FT /tag- a

EP-304013-A.
 PD 22-FEB-1989.
 PF 16-AUG-1988; 113283.
 PR 15-AUG-1987; JP-203298.
 PR 21-AUG-1987; JP-207995.
 PR 18-NOV-1987; JP-291404.
 PR 11-DEC-1987; JP-313228.
 PR 05-FEB-1988; JP-025260.
 PR 10-FEB-1988; JP-029366.
 PR 19-FEB-1988; JP-037905.
 PR 25-MAY-1988; JP-125560.
 PA (ASAH) Asahi Kasei Kogyo.
 PI Kitaguchi N, Takahashi Y, Tokushima Y, Itoh H;
 DR WPI: 89-055458/08.
 PT Human senile plaque amyloid precursor protein and DNA -
 PT used for study and diagnosis of dysbolism in the central nervous
 PT system
 PS Disclosure; pp; English.
 CC The gene product INS76, which is claimed as a senile plaque amyloid
 CC precursor (SPAP), has been shown to comprise four exons H, I, J and K.
 CC The SPAP is useful for diagnosis of dysbolism in the CNS such as senile
 CC dementia.
 SQ Sequence 953 BP; 252 A; 188 C; 186 G; 327 T;

Query Match 100.0%; Score 9; DB 1; Length 953;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 522 ttctcagaa 530
 QY 25 ttctcagaa 33

RESULT 162

ID T96064 standard; cDNA; 958 BP.
 AC T96064.
 DT 08-JUN-1998 (first entry)
 DE Apoptosis associated protein bdk gene.
 KW Apoptosis associated protein; bdk gene; human; cell death;
 KW cell proliferation; degenerative disorder; gene therapy;
 KW Alzheimer's disease; autoimmune disease; cancer; inflammation; ds.
 OS Homo sapiens.
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 51..800
 FT /tag- a
 FT variation 157..158
 FT /tag- b
 FT /note- "AAG insert in some sequenced clones"

W09745128-A1.
 PD 04-DEC-1997.
 PE 29-MAY-1997; U09194.
 PR 29-MAY-1996; US-632514.
 PA (APOP-) APOPTOSIS TECHNOLOGY INC.
 PI Gallo GJ;
 DR WPI: 98-032326/03.
 DR P-PDB: W8358.
 PT Human apoptosis associated protein Bdk - useful to treat

PT degenerative disorders characterised by inappropriate cell death or proliferation
 PS Claim 11: Fig 2, 110pp: English.
 CC This nucleotide sequence includes a coding region for novel human
 CC apoptosis associated protein Bbk (see W8358), a member of the
 CC Bcl-2 family that can induce apoptosis in cells and oppose the
 CC function of Bcl-2 and related cell death suppressors in cells. The
 CC Bbk gene was identified using a yeast two-hybrid system in which a
 CC GAL4 DNA binding domain/Bak fusion protein was used as the bait to
 CC isolate Bak-interacting proteins expressed as GAL4 activation
 CC domain fusions generated from a cDNA library of Epstein-Barr virus
 CC transformed human B lymphocytes. The Bbk protein and its BH3 domain
 CC peptide (see W8346) can be used to induce apoptosis in non-apoptotic
 CC cells, useful in the treatment of degenerative disorders
 CC characterised by inappropriate cell death, e.g. Alzheimer's disease,
 CC or inappropriate cell proliferation, e.g. cancer, autoimmune disease,
 CC tissue hypertrophy and inflammatory disorders. They may be
 CC administered with a carrier, e.g. tissue specific receptor, or
 CC preferably via a viral vector or liposome, in the form of an
 CC encoding nucleic acid sequence. Vectors and host cells are
 CC claimed, as well as a labelled nucleotide probe that can be used
 CC e.g. to monitor the progress of degenerative disorders associated
 CC with the expression of Bbk.
 SQ Sequence 958 BP: 163 A; 314 C; 282 G; 199 T;
 Query Match 100.0%; Score 9; DB 39; Length 958;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 806 ttctcagaa 814
 |||||||
 Cp 33 ttctcagaa 25

RESULT 163
 ID Q79929 standard; cDNA to mRNA; 961 BP.
 AC Q79929;
 DT 01-SEP-1995 (first entry)
 DE Anti-tobacco mosaic virus monoclonal Ab light chain cDNA.
 KW Tobacco mosaic virus; TMV; monoclonal antibody;
 KW light chain; virus-resistant plants; biofarming; ss.
 FS Synthetic.
 FH Key Location/Qualifiers
 FT cds 34..753
 FT /*tag- a
 FT signal_peptide 34..93
 FT /*tag- b
 FT mat_peptide 94..750
 FT /*tag- c
 PN J06319396-A.
 PD 22-NOV-1994.
 PF 07-MAY-1993; 131208.
 PR 07-MAY-1993; JP-131208.
 PA (NISB) JAPAN TOBACCO INC.
 PA (KURS) KURARAY CO LTD.
 DR WPI: 95-040220/06.
 DR P-PSDB: R66757.
 PT Transformed plant producing animal-derived anti-virus antibody -
 PT esp. tobacco plants producing anti-tobacco mosaic virus
 PT monoclonal antibody
 PS Example 2; Pages 13-14; 26pp; Japanese.
 CC Q79929 and Q79930 encode R66757 and R66758, the light and heavy
 CC chains of an animal derived anti-tobacco mosaic virus (TMV)
 CC monoclonal antibody. The cDNAs were incorporated into a Ti
 CC plasmid vector, which was incorporated into A. tumefaciens.
 CC The resultant plant expression vector was used to transform
 CC tobacco plants, making them TMV resistant, the plants could
 CC also be biofarmed for the prodn. of anti-virus antibodies.
 SQ Sequence 961 BP: 243 A; 255 C; 213 G; 250 T;

Query Match 100.0%; Score 9; DB 14; Length 961;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 ttctcagaa 34
 |||||||
 Cp 33 ttctcagaa 25

RESULT 164
 ID T68932 standard; DNA; 972 BP.
 AC T68932;
 DT 23-JUL-1997 (first entry)
 DE Bovine PIT-1 exon 6 sequence.
 KW Primer; polymerase chain reaction; PCR; amplify; detection; RFLP;
 KW HinfI polymorphism; exon 6; bovine; pituitary; PIT-1 gene; genotype;
 KW high milk production; restriction fragment length polymorphism; ss.
 OS Bos taurus.
 FH Key Location/Qualifiers
 FT variation 630
 FT /*tag- a
 FT /note- "Position of HinfI polymorphism"
 FT exon 468..678
 FT /*tag- b
 FT /note- "Encodes the last 70 amino acids of PIT-1"
 FT complement (424..443)
 FT /*tag- c
 FT /note- "Binds primer given in T68930"
 FT primer_bind 855..874
 FT /*tag- d
 FT /note- "Binds primer given in T68931"
 PN U55614364-A.
 PD 25-MAR-1997.
 PF 16-MAY-1994; 243543.
 PR 16-MAY-1994; US-243543.
 PA (IOWA) UNIV IOWA STATE RES FOUND INC.
 PI Freeman AE, Tuggle CK;
 DR WPI: 97-201471/18.
 DR P-PSDB: W18344.
 PT Identification of cows giving high milk yields - on basis of
 PT polymorphism in exon 6 of PIT-1 gene
 PS Claim 6: Fig 4; 22pp; English.
 CC This sequence represents exon 6 of the bovine pituitary derived PIT-1
 CC gene, showing the HinfI polymorphic site. This sequence was amplified
 CC using the primers given in T68930-31 for detecting the HinfI
 CC polymorphism. The method of the invention identifies a bovine with
 CC a genotype indicative of high milk production. The method comprises
 CC obtaining a nucleic acid sample from the bovine and ascertaining that
 CC the base at nucleotide position 165 in exon 6 of the bovine PIT-1 gene
 CC is adenine, or obtaining a sample of genomic DNA, digesting the sample
 CC with HinfI, separating the resulting fragments, and ascertaining that
 CC there is no HinfI site at nucleotide position 165 in exon 6 of the bovine
 CC PIT-1 gene. These primers are used to amplify the bovine PIT-1 gene
 CC sequence, then restriction fragment length polymorphism analysis is
 CC performed to detect an 800 bp fragment [indicating absence of a HinfI
 CC site]. Cattle with the genotype are selected for breeding. The genotype
 CC lacking the HinfI site is associated with an average 13.2% increase in
 CC milk production and an average 17.2% increase in total protein.
 SQ Sequence 972 BP: 287 A; 190 C; 176 G; 315 T;

Query Match 100.0%; Score 9; DB 30; Length 972;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 851 ttctcagaa 859
 |||||||
 Cp 33 ttctcagaa 25

RESULT 165
 ID V59982 standard; cDNA to mRNA; 987 BP.
 AC V59982;
 DT 25-NOV-1998 (first entry)
 DE Nucleic acid rc2 encoding human homologue of rat acyl-CoA synthetase.
 KW Human homologue; yeast Rad50; Drosophila Septin-2; Acyl-CoA synthetase;
 KW immunomodulatory activity; identification; activated T-cell; ds.

OS - Homo sapiens.
 PN M09838306-A1.
 PD 03-SEP-1998.
 PF 27-FEB-1997; U03159.
 PR 27-FEB-1997; WO-003159.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Dolganov G;
 DR WPI: 98-481207/41.
 PT Novel human immunomodulatory poly(peptide(s) - have homology to the
 PS yeast RAD50 or Drosophila Septin-2 proteins
 SS Disclosure: Page 64; 15pp; English.
 CC V59381-89 represent multiple isoforms of a human homologue of the rat
 CC brain long chain acyl-CoA synthetase. The present sequence has 90%
 CC overall homology to the rat gene, and is expressed in activated
 CC T-cells, testis, and foetal liver tissues. The specification
 CC also describes sequences encoding human homologues of the
 CC yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
 CC proteins have immunomodulatory activity. The nucleic acids and
 CC proteins can be used to identify activated T-cells in a sample
 CC population. They can also be used to isolate and identify sequences
 CC encoding other proteins or other compounds having immunomodulatory
 CC activity.
 SQ Sequence 987 BP; 267 A; 223 C; 223 G; 274 T;
 Query Match 100.0%; Score 9; DB 49; Length 987;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 9 ttctgagaa 17
 |||||
 QY 25 ttctgagaa 33
 RESULT 166
 ID V24131 standard; DNA: 993 BP.
 AC V24131;
 DT 28-SEP-1998 (first entry)
 DE Homo sapiens TCl52 sequence.
 KW BARD1, ring protein; BRCA1; breast cancer; risk; diagnosis; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..993
 FT /*tag= a
 FT /product= TCl52 protein
 FT
 PN M09812327-A2.
 PD 26-MAR-1998.
 PF 19-SEP-1997; U16842.
 PR 04-APR-1997; US-042985.
 PR 20-SEP-1996; US-025296.
 PR 03-APR-1997; US-042611.
 PA (TEKA) UNIV TEXAS SYSTEM.
 PI Baer R, Bowcock AM;
 DR WPI: 98-230317/20.
 DR P-PSDB; V54095.
 PT DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
 PT as breast cancer antigen, BRCA1, binding proteins are useful to
 PT identify patient having or at risk of developing cancer
 PS Claim 25; Page 200-201; 34pp; English.
 CC The sequence is that encoding a protein which can be used in the
 CC preparation of the recombinant breast cancer antigen, BRCA1, binding
 CC proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the
 CC detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
 CC sequence, specifically a wild type BARD1 composition for the detection
 CC or purification of BRCA1, useful to identify a patient having, or at
 CC risk of developing cancer. BARD1 can be used in the preparation of an
 CC anti-BARD1 antibody, and in the detection and purification of a BRCA1
 CC protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
 CC identification of a binding protein agonist or antagonist that alters
 CC the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the
 CC biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445
 CC complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
 CC BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
 CC a patient having or at risk of developing cancer.

SQ Sequence 993 BP; 386 A; 184 C; 184 G; 239 T;
 Query Match 100.0%; Score 9; DB 45; Length 993;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 762 ttctgagaa 770
 |||||
 QY 25 ttctgagaa 33
 RESULT 167
 ID T51349 standard; DNA: 1016 BP.
 AC T51349;
 DT 29-APR-1997 (first entry)
 DE Bacillus subtilis protein secretion chaperone ftsy gene.
 KW Protein secretion chaperone; ftsy gene; ds.
 OS Bacillus subtilis strain 168.
 FH Key Location/Qualifiers
 FT cds 19..1008
 FT /*tag= a
 FT /product= Ftsy
 FT
 PN M09703197-A1.
 PD 30-JAN-1997.
 PF 05-JUL-1996; NL0278.
 PR 07-JUL-1995; EP-201871.
 PA (GENV) GENENCOR INT BV.
 PI Broekhuizen CP, Kerkman R, Quax WJ;
 DR WPI: 97-119052/11.
 DR P-PSDB; W11328.
 PT Secretion factor from Gram positive bacteria encoded by the ftsy
 PT gene - used to increase secretion of proteins being overexpressed by
 PT cells
 PS Claim 1; Page 28-30; 51pp; English.
 CC The Bacillus subtilis ftsy gene (T51349) codes for protein secretion
 CC chaperone Ftsy (W11328), which can be used to increase secretion of
 CC overexpressed proteins (homologous or heterologous) from Gram-
 CC positive (esp. Bacillus) hosts. To isolate the gene, degenerate
 CC primers (see also T51331-53) were synthesized making use of existing
 CC homology boxes between the SR-alpha homologues of different
 CC organisms. After an inverted PCR using a 110 bp fragment, which was
 CC derived from a nested PCR, a 4 kb fragment (see also S1356) was
 CC detected that included the ftsy coding sequence. The gene can be
 CC incorporated into a recombinant vector pref. also contg. regulatory
 CC elements and opt. the gene encoding the protein of interest. By
 CC leading the overexpressed protein into the secretion pathway, the
 CC chaperone facilitates purificn., may provide higher yields, prevents
 CC protein aggregation and makes possible continuous culture and prodn.
 SQ Sequence 1016 BP; 346 A; 182 C; 252 G; 236 T;
 Query Match 100.0%; Score 9; DB 28; Length 1016;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 303 ttctgagaa 311
 |||||
 QY 25 ttctgagaa 33
 RESULT 168
 ID V69789 standard; DNA: 1021 BP.
 AC V69789;
 DT 05-FEB-1999 (first entry)
 DE CTLA4-ova fusion protein encoding DNA.
 KW CTLA4 receptor; CTLA4-1g; fusion protein; B7 antigen; hinge; CH2; CH3;
 KW extracellular domain; human; immunoglobulin; T cell; immune system;
 KW autoimmune disease; cancer; viral infection; ovalbumin; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT sig_peptide 1..77
 FT /*tag= a
 FT /note= "CDS leader sequence"
 FT misc_feature 78..448

FT /*tag- b
FT /note- "sequence encoding CT1A4 amino-terminal domain"
FT misc.feature 449..1021
FT /*tag- c
FT /note- "sequence encoding ovalbumin carboxy-terminal domain"

US5844095-A.
PD 01-DEC-1998.
PF 18-JAN-1995: 375390.
PR 18-JAN-1995: US-375390.
PR 27-JUN-1991: US-723617.
PR 22-JAN-1993: US-008898.
PR 28-MAY-1993: US-069693.
PR 15-APR-1994: US-228208.
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Brady W, Damlie NK, Ledbetter JA, Linsley PS;
PI WPI: 99-044666/04.
PT Fusion protein of CTLA4 and immunoglobulin fragment - for treating immune system disorders
PS Example 8: Fig 26: 75pp: English.
CC This represents a nucleic acid sequence encoding a CTLA4-ova fusion protein. The invention provides a CTLA4-19 fusion protein that binds the CC B7 antigen and has a first amino acid sequence consisting of the CC extracellular domain of CTLA4 and a second amino acid sequence consisting of the hinge, CH2 and CH3 regions of a human immunoglobulin molecule. The CC fusion protein inhibits interaction of T cells with B7-positive cells and CC may be useful for treating immune system diseases, e.g. autoimmune diseases, cancer or viral infections. The fusion protein encoded by the CC present sequence is an example of a soluble fusion protein and contains CC an amino-terminal CTLA4 domain and an ovalbumin carboxy-terminal domain.
SQ Sequence 1021 BP: 260 A; 246 C; 270 G; 245 T;

Query Match 100.0%; Score 9; DB 54; Length 1021;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 529 ttctcagaa 537
 |||||
Qy 25 ttctcagaa 33

RESULT 169
ID T13161 standard; DNA: 1022 BP.
AC T13161;
DE 23-MAY-1996 (first entry)
DE SCSV segment 2.
KW SCSV; promoter; transcription; transgenic plant; legume;
KW gene expression; crop improvement; ss.
OS Subterranean clover stunt virus isolate F.
FH Key Location/Qualifiers
FT cds 82..924
FT /*tag- a

W09606932-A1.
PD 07-MAR-1996.
PF 30-AUG-1995: AU0552.
PR 30-AUG-1994: AU-007770.
PR 07-NOV-1994: AU-009281.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PI (AUSU) UNIV AUSTRALIAN NAT.
PI Boevink PC, Chu PNG, Keese PK, Khan RI, Larkin PJ;
PI Marshall JS, Surin BP, Taylor WC, Waterhouse PM;
PI WPI: 96-160363/16.
PT Circovirus transcription regulatory sequences and related constructs
PT - useful in plants, esp. leguminous plants, for the modulation of gene expression
PS Claim 4: Page 56: 121pp: English.
CC The subterranean clover stunt virus (SCSV) genome has at least 7 CC distinct ssDNA components, designated segments 1-7 (T13160-66), each CC contg. 1 major open reading frame and a non-coding region. Segment 2 is predicted to be a viral replication-associated protein gene.
CC Genetic constructs useful in the genetic engineering of plants (esp. CC legumes) comprise 1 or more heterologous gene(s) operatively linked to a promoter region, and in some cases also a terminator region.

CC selected from segments 1-7. The transcription regulators facilitate CC expression of foreign genes in plants and also facilitate control of CC levels of gene expression in different plant tissue types.
SQ Sequence 1022 BP: 309 A; 183 C; 245 G; 285 T;

Query Match 100.0%; Score 9; DB 18; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 315 ttctcagaa 323
 |||||
Cp 33 ttctcagaa 25

RESULT 170
ID V23014 standard; DNA: 1026 BP.
AC V23014;
DE 05-AUG-1998 (first entry)
DE DNA encoding tryptophanyl tRNA synthetase of Streptococcus pneumoniae.
KW Tryptophanyl tRNA synthetase; trps; Streptococcus pneumoniae sp. 0100993;
KW Clostridium longisporum; antagonist; treatment; inhibition; vaccine; ss.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT cds 1..1026
FT /*tag- a

W09810652-A1.
PD 19-MAR-1998.
PF 12-SEP-1997: U16367.
PR 12-SEP-1996: GB-019072.
PA (SMIR) SMITHKLINE BEECHAM CORP.
PI (SMIR) SMITHKLINE BEECHAM PLC.
PI Gentry DR, Greenwood RC, Lawlor EJ;
PI WPI: 98-207073/18.
DR P-RSDB: W56423.
DR Streptococcus pneumoniae tryptophanyl tRNA synthetase - useful to PT identify drugs and raise antibodies for protection and diagnosis of PT Streptococcal infection
PS Claim 4: Pages 8-9; 43pp: English.
CC The present sequence encodes a tryptophanyl tRNA synthetase (trps) of CC Streptococcus pneumoniae sp. 0100993. This trps is structurally related to other proteins of the tryptophanyl tRNA synthetase family. The protein CC exhibits greatest homology to Clostridium longisporum tryptophanyl tRNA CC synthetase protein (81% similarity). Compounds which interact with and CC inhibit or activate trps can be identified by competitive binding assays CC involving contact with the trps protein. Antagonists can be used for CC treatment where inhibition of trps is required. Disease related to trps CC can be determined by determining the presence of nucleic acids encoding trps in samples. trps can also be used to generate an immunological CC response in a mammal, by inoculation of trps, adequate to produce CC antibody and/or T cell immune response to protect against disease.
SQ Sequence 1026 BP: 323 A; 159 C; 239 G; 305 T;

Query Match 100.0%; Score 9; DB 42; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 351 ttctcagaa 359
 |||||
Cp 33 ttctcagaa 25

RESULT 171
ID V68168 standard; DNA: 1039 BP.
AC V68168;
DE 15-FEB-1999 (first entry)
DE Porcine Interferon-beta 1 cDNA.
KW Interferon-beta 1; fibroblast interferon; virucide; antiviral;
KW antitumour; pig; ss.
OS Sus scrofa.
FH Key Location/Qualifiers
FT cds 200..763
FT /*tag- a

sig_peptide 200..266


```
FT      mat_peptide      /*tag= b
FT      269..760
FT      /*tag= c
PN      US5831023-A.
PD      03-NOV-1998.
PF      19-MAY-1995: 444454.
PR      02-OCT-1987: US-104461.
PR      08-MAR-1982: US-355298.
PR      01-NOV-1982: US-438128.
PR      23-AUG-1991: US-749371.
PR      21-SEP-1992: US-949327.
PR      19-MAY-1995: US-444454.
PA      (GETH ) GENENTECH INC.
PI      Capon DJ, Goeddel DV.
DR      P-PSDB: W83903.
PT      Animal interferon polypeptides - useful as antiviral and antitumour
PT      agents
PS      Disclosure: Fig 14B: 48bp; English.
CC      This cDNA sequence includes an open reading frame encoding porcine
CC      interferon beta-1 (see W83903). The invention provides bovine,
CC      murine, porcine, feline, laprine and rat mature alpha, beta and
CC      gamma interferons and their propeptides (see W83895-905, W83918 and
CC      W70582-85). Recombinant DNA methods can be utilised to prepare
CC      these non-human animal interferons in amounts sufficient to enable
CC      the determination of their biochemical properties and bioactivities,
CC      and hence to allow efficient production for commercial or biological
CC      exploitation. Interferons are useful as antiviral and antitumour
CC      agents.
SO      Sequence 1039 BP; 299 A; 219 C; 242 G; 279 T;

Query Match      100.0%; Score 9; DB 54; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      468 ttccagaa 476
Cp      33 ttccagaa 25

RESULT 172
ID      V08185 standard; cDNA; 1039 BP.
AC      V08185.
DE      Porcine interferon-beta 1 DNA.
KW      Interferon-beta 1; fibroblast interferon; virucide; antiviral;
KW      antitumour; pig; ss.
OS      Sus scrofa.
FH      Key Location/Qualifiers
FT      CDS 200..763
FT      /*tag= a
FT      sig_peptide 200..268
FT      /*tag= b
FT      mat_peptide 269..760
FT      /*tag= c
PN      US5827694-A.
PD      27-OCT-1998.
PF      19-MAY-1995: 446171.
PR      02-OCT-1987: US-104461.
PR      08-MAR-1982: US-395298.
PR      01-NOV-1982: US-438128.
PR      23-AUG-1991: US-749371.
PR      21-SEP-1992: US-949327.
PR      19-MAY-1995: US-446171.
PA      (GETH ) GENENTECH INC.
PI      Capon DJ, Goeddel DV.
DR      P-PSDB: W73232.
PT      DNA encoding non-human interferon polypeptide(s) - useful for
PT      producing recombinant polypeptide(s) to determine their
PT      bioactivity
PS      Claim 1: Fig 14B: 59bp; English.
CC      This cDNA sequence includes an open reading frame encoding porcine
```

```
CC      interferon beta-1 (see W73232). The invention provides bovine,
CC      porcine, feline and rabbit mature interferons and their propeptides
CC      (see W73224-35). Recombinant DNA methods can be used to prepare
CC      these non-human animal interferons in amounts sufficient to enable
CC      the determination of their biochemical properties and
CC      bioactivities, and hence to allow efficient production for
CC      commercial or biological exploitation. Interferons are useful as
CC      antiviral and antitumour agents. A replicable expression vehicle
CC      comprising claimed non-human animal interferon nucleic acid, a
CC      process for producing such an interferon, and a microorganism
CC      (especially Escherichia coli) or cell culture transfected with the
CC      nucleic acid are claimed.
SO      Sequence 1039 BP; 299 A; 219 C; 242 G; 279 T;

Query Match      100.0%; Score 9; DB 54; Length 1039;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      468 ttccagaa 476
Cp      33 ttccagaa 25

RESULT 173
ID      V04889 standard; DNA; 1044 BP.
AC      V04889.
DE      30-APR-1998 (first entry)
DE      Nucleic acid sequence of the specification.
KW      Neo-minichromosome; multicentric; dicentric; chromosome; euchromatin;
KW      heterochromatic DNA; megachromosome; artificial chromosome;
KW      selectable marker; satellite artificial chromosome; SATC; genomic locus;
KW      targeted integration; transgenic animal; therapeutic product;
KW      gene therapy; cloning vehicle; genomic DNA library;
KW      chromosome fragmentation; ss.
OS      unidentified.
PN      WO9740183-A2.
PD      30-OCT-1997.
PF      10-APR-1997: 005911.
PR      07-AUG-1996: US-695191.
PR      10-APR-1996: US-628822.
PR      15-JUL-1996: US-682080.
PA      (AMGE-) AMERICAN GENE THERAPY INC.
PA      (BIOL-) BIOLOGICAL RES CENT HUNGARIAN ACAD SCI.
PA      (DVILO-) UNIV IOMIA LINDA.
PI      Hadlaczky G, Szalay AA.
DR      WPI: 97-353860/49.
PT      Producing satellite artificial chromosomes or mini-chromosomes -
PT      useful for, e.g. cloning multiple proteins of a metabolic pathway or
PT      multivalent vaccines, etc.
PS      Example 12: Page 197; 248bp; English.
CC      The present sequence appears in the specification, but no further
CC      information is provided. The specification describes the production of
CC      artificial chromosomes. Artificial chromosomes can be produced by
CC      incorporating a DNA fragment comprising a selectable marker into a
CC      cells' genomic DNA, into or adjacent to an amplifiable region, and
CC      selecting a cell that comprises either a satellite artificial
CC      chromosome (SATAC) (e.g. V04903-09), a de novo centromere, or a
CC      minichromosome (e.g. V04900-02). Artificial chromosomes provide an extra
CC      genomic locus for targeted integration of megabase size DNA fragments
CC      that contain single or multiple genes. SATACs can be introduced into
CC      embryonic cells of non-human animals to produce transgenic animals that
CC      express a heterologous DNA encoding a therapeutic product, e.g. anti-HIV.
CC      The DNA may also encode antigens that induce immunoprotective response
CC      against pathogen. These therapeutic products can be used in gene therapy.
CC      The artificial chromosomes are useful as cloning vehicles that
CC      accommodate entire genomes for the preparation of genomic DNA libraries,
CC      and also for the production of proteins which may be involved in a
CC      biochemical pathway or in multivalent vaccines.
SO      Sequence 1044 BP; 299 A; 184 C; 232 G; 329 T;

Query Match      100.0%; Score 9; DB 38; Length 1044;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 52 ttctcagaa 60
 |||||
 Cp 33 ttctcagaa 25

RESULT 174
 ID 188112 standard; DNA: 1047 BP.
 AC 188112;
 DT 28-APR-1998 (first entry)
 DE Phenylalanyl tRNA synthetase alpha subunit encoding DNA.
 KM Phenylalanyl tRNA synthetase: phes protein; bacterial infection;
 KW meningitis; vaccine; antibacterial; immune response; ss.
 OS Streptococcus pneumoniae.
 FH Key Location/Qualifiers
 FT CDS 1..1047
 FT /tag- a
 FT /product= "phes alpha subunit"

PN W09738722-A1.
 PD 23-OCT-1997.
 PF 18-APR-1997: U06876.
 PR 18-APR-1996: GB-007993.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Lawlor EJ;
 PI WPI: 97-526215/48.
 DR P-PSDB: W27668.
 DR Polynucleotide encoding Streptococcus pneumoniae phenylalanyl tRNA
 PT synthetase alpha and beta subunits - useful to diagnose, treat or
 PT prevent bacterial infection, especially meningitis
 PS Claim 5: Page 45: 53pp; English.
 CC The present sequence encodes the alpha subunit of phenylalanyl tRNA
 CC synthetase (phes) from Streptococcus pneumoniae 0100993 (NCIMB 40794).
 CC The polynucleotide encoding phes can be used for the recombinant
 CC production of phes, which can be used to treat conditions requiring phes
 CC activity. Sequences antisense to the polynucleotide can be used to
 CC control phes expression. Phes, or phes expressing vectors can be used to
 CC induce an immune response, i.e. an antibody (Ab) and/or T cell response,
 CC against S. pneumoniae to protect against infection, or to screen for
 CC antagonists or agonists of the polynucleotide encoding phes or phes's
 CC activity, i.e. antibacterials. The antagonist, e.g. an anti-phes Ab, can
 CC be used to treat conditions requiring phes inhibition, e.g. an
 CC S. pneumoniae infection, particularly meningitis. A fragment of the
 CC polynucleotide encoding phes can be used as a probe to isolate full length
 CC or related sequences, or diagnose, e.g. by polymerase chain reaction.
 CC Infection stage and type, including mutation and polymorphism detection.
 CC Diagnosis may also be achieved by detecting phes gene overexpression,
 CC e.g. by immunosay. The Ab can be used to treat infection, isolate or
 CC identify phes expressing clones, purify phes and as an immunoassay
 CC reagent. More generally, the products can prevent adhesion of bacteria
 CC to wounds and in dwelling devices, block phes protein mediated invasion
 CC of mammalian cells and block the normal progression of infection.
 CC Sequence 1047 BP: 284 A; 221 C; 261 G; 281 T;

Query Match 100.0%; Score 9; DB 37; Length 1047;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1027 ttctcagaa 1035
 |||||
 Cp 33 ttctcagaa 25

RESULT 175
 ID 057986 standard; DNA: 1060 BP.
 AC 057986;
 DT 20-AUG-1994 (first entry)
 DE Genomic sequence of human CD8.
 KM gp39: T-cell antigen; CD40 ligand; B-cell proliferation;
 KW CD8: fusion protein; CD8: fusion protein; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 66..773

FT EP-585943-A. /tag- a
 PD 09-MAR-1994.
 PF 03-SEP-1993: 114153.
 PR 04-SEP-1992: US-940605.
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Aruffo AA, Hollenbaugh D, Ledbetter JA, Aruffo A;
 DR WPI: 94-076264/10.
 DR P-PSDB: R59550.
 PT New nucleic acid encoding human gp39 T cell antigen - which is a
 PT ligand for the CD40 receptor, causing proliferation and
 PT differentiation of B cells and some cancer cells
 PS Disclosure: Fig 9: 39pp; English.
 CC The complete nucleic acid sequence of human gp39 (hg39) protein
 CC (corresp. to cDNA) and the complete AA sequence of hg39 are
 CC presented in Q57984 and R49548 respectively and contd. In plasmid
 CC CD8-hg39, deposited with the ATCC as E. coli, CD8 MC1061/p3-hg39
 CC and assigned accession No. 69050. The human T cell antigen gp39 is a
 CC ligand for the CD40 receptor. Soluble gp39 may be produced using the
 CC expression vector CD8-gp39. Chimeric genes may be constructed by
 CC fusing sequences encoding the extracellular domains of gp39 and
 CC CD8, pref. murine or human CD8 protein. Plasmid p3-shg39 encoding
 CC the fusion protein of the extracellular domains from gp39 and CD8
 CC is deposited in E.coli as ATCC 69049.
 CC Sequence 1060 BP: 196 A; 349 C; 294 G; 221 T;

Query Match 100.0%; Score 9; DB 10; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1049 ttctcagaa 1057
 |||||
 Cp 33 ttctcagaa 25

RESULT 176
 ID V30311 standard; DNA: 1073 BP.
 AC V30311;
 DT 28-SEP-1998 (first entry)
 DE Bacillus thuringiensis MIS-6 insecticidal toxin 196J4 DNA.
 KM Insecticidal; pesticide; toxin; MIS-6; delta-endotoxin;
 KW biological control; lepidopteran; coleopteran; ss.
 OS Bacillus thuringiensis strain PS196J4 (NRRL B-21860).
 FH Key Location/Qualifiers
 FT CDS 2..1072
 FT /tag- a

PN W09818932-A2.
 PD 07-MAY-1998.
 PF 30-OCT-1997: U19804.
 PR 30-OCT-1996: US-029848.
 PA (MYCO) MYCOGEN CORP.
 PI Dullum CJ, Feltelson JS, Loewer D, Muller-Cohn J,
 PI Narva KE, Schmeits JL, Schnepf HE, Schwab G, Stamp L,
 PI Stockhoff BA;
 DR WPI: 98-272226/24.
 DR P-PSDB: W60228.
 PT Bacillus thuringiensis isolates - used for producing pesticidal
 PT toxins and nucleotide sequences for control of lepidopterans and
 PT coleopterans
 PS Claim 5: Page 92-93; 139pp; English.
 CC This DNA sequence encodes a novel soluble toxin (see W60228) of
 CC Bacillus thuringiensis (B.t.) strain PS196J4 (NRRL B-21860). This
 CC toxin belongs to the novel MIS-6 family of B.t. toxins that have
 CC toxicity against non-mammalian pests. The novel DNA was obtained
 CC by PCR amplification (see V30298-99) of total cellular genomic
 CC DNA. It can be used to produce recombinant hosts (preferably plant
 CC or bacterial) that express the toxin, or as a PCR primer or
 CC hybridisation probe for use in identifying and characterising
 CC MIS-6 family toxin genes. Disclosed and claimed are novel B.t.
 CC isolates and toxins (see W60218-32) which have activity against
 CC lepidopteran and/or coleopteran pests. Isolated genes, probes
 CC and primers (see V30288-321 and 199734-87), and transformed host
 CC cells. The invention provides 8 entirely new families of toxins.

CC including MIS-6, from B.t. isolates. MIS toxins are also useful
CC for their ability to form pores in cell membranes, and can be used
CC to facilitate entry of a second agent into a target cell.

SO Sequence 1073 BP; 378 A; 193 C; 218 G; 284 T;

Query Match 100.0%; Score 9; DB 44; Length 1073;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 811 tctcgagaa 819
|||||

OY 25 tctcgagaa 33

RESULT 177
ID T98694 standard; DNA; 1087 BP.

AC T98694;

DE DNA encoding a S. pneumoniae anthranilate synthase component.

DE Streptococcus pneumoniae protein; genetic immunisation; antagonist;

KW Immunological response; inoculation; antibody production; inhibitor;

KW T cell immune response; antimicrobial compound; bacterial adhesion;

KW extracellular matrix protein; protein-mediated cell invasion; wound;

KW pathogenesis; ss.

OS Streptococcus pneumoniae.

PH Key Location/Qualifiers

FT CDS 117..683

FT WO9743303-A1. /*tag= a

PN 20-NOV-1997.

PD 14-MAY-1997; U07950.

PF 14-MAY-1996; US-017670.

PR (SMK) SMITHKLINE BEECHAM CORP.

PA (SMK) SMITHKLINE BEECHAM PLC.

PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,

PI Stodola RK;

PI WPI: 98-008793/01.

DR P-PSDB: W38648.

PT Novel Streptococcus pneumoniae proteins and related DNA - useful for

PT diagnosing anti-microbial agents for treatment of bacterial

PT infections

PS Claim 4; Pages 201-202; 483pp; English.

CC This sequence encodes a Streptococcus pneumoniae protein that (based on

CC homology with a Lactococcus lactis protein) is an anthranilate synthase

CC component, and represents a DNA sequence of the invention.

CC The DNA sequences were isolated from Streptococcus pneumoniae strain

CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the

CC invention can be used to identify compounds which interact with and

CC inhibit or activate the activity of the proteins. Antagonists can be

CC used to treat diseases caused by S. pneumoniae proteins, through genetic

CC immunisation. They can also be used to induce an immunological response

CC in a mammal by inoculation with the S. pneumoniae proteins or delivery

CC and/or T cell immune responses to protect the animal from disease. The

CC proteins can also be used to identify antimicrobial compounds which are

CC capable of inhibiting their bioactivity. In particular the proteins of

CC extracellular matrix proteins on in-dwelling devices or in wounds, to

CC block protein-mediated mammalian cell invasion, and to block the normal

CC progression of pathogenesis in infections initiated other than by the

CC implantation of in-dwelling devices or other surgical techniques.

SO Sequence 1087 BP; 314 A; 208 C; 264 G; 299 T;

Query Match 100.0%; Score 9; DB 47; Length 1087;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 196 tctcgagaa 204
|||||

OY 25 tctcgagaa 33

RESULT 178

ID T91592 standard; cDNA; 1092 BP.

AC T91592;

DE 20-APR-1998 (first entry)

DE Rat UBCE2A coding sequence.

KW Ubiquitin conjugating enzyme; UBCE2A; rat; ubiquitin; cell proliferation;

OS Rattus sp.

PH Key Location/Qualifiers

FT CDS 82..558

FT WO9735961-A1. /*tag= a

PN 02-OCT-1997.

PD 28-MAR-1997; U05337.

PF 28-MAR-1996; US-014388.

PR (HARD) HARVARD COLLEGE.

PA Haber E, Kuo C, Lee M;

PI WPI: 97-489625/45.

DR P-PSDB: W32487.

PT Mammalian UBCE2A catalyses covalent attachment of ubiquitin to E2A -

PT inhibitors of which can be used to inhibit proliferation of cell

PT that express transcription factor E2A

PS Claim 7; Page 47; 69pp; English.

CC This sequence represents the coding sequence for the rat UBCE2A

CC (ubiquitin conjugating enzyme) of the invention. The encoded protein

CC catalyses the covalent attachment of ubiquitin to E2A. An anti-UBCE2A

CC antibody or a single stranded nucleic acid that is antisense to a portion

CC of this sequence that reduces the level of UBCE2A biological activity in

CC a cell can be used to inhibit the proliferation of a cell that expresses

CC transcribed factor E2A. The proliferation of a cell may also be

CC inhibited by introducing a proteasome inhibitor into the cell or by

CC introducing into the cell a mutant E2A that possesses the transcription

CC activator activity of wild type E2A but lacks the UBCE2A binding site or

CC lacks at least 1 of the lysine residues which are ubiquitination sites on

CC the wild type E2A.

SO Sequence 1092 BP; 286 A; 264 C; 259 G; 283 T;

Query Match 100.0%; Score 9; DB 38; Length 1092;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 935 tctcgagaa 943
|||||

Cp 33 tctcgagaa 25

RESULT 179
ID V44860 standard; cDNA; 1097 BP.

AC V44860;

DE 21-OCT-1998 (first entry)

DE Clone Bg142.1 coding sequence.

KW Secreted protein; nutritional source; cell proliferation activity;

KW cell differentiation activity; immune stimulant; tissue growth activator;

KW haematopoiesis regulator; anti-inflammatory; tumour invasion suppressor;

KW tumour inhibitor; Clone Bg142.1; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 62..760

FT WO9825962-A2. /*tag= a

PN 18-JUN-1998.

PD 12-DEC-1997; U23224.

PF 11-DEC-1997; US-989232.

PR 13-DEC-1996; US-766263.

PA (GENT) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D,

PI Racie LA, Spaulding V, Treacy M;

PI WPI: 98-362424/31.

DR P-PSDB: W69241.

PT New isolated polynucleotides - obtained from human adult testis,

PT human adult ovary, human adult brain and human adult heart cDNA

PT libraries

PS Claim 17; Page 68; 108pp; English.

CC This sequence represents a polynucleotide of the invention, and encodes a

CC secreted protein. It was isolated from a human adult brain cDNA library,
CC and is designated clone B6142.1. The DNA sequences and encoded
CC polypeptides can be used as nutritional sources or supplements, or may
CC exhibit e.g. cytokine and cell proliferation/differentiation activity,
CC immune stimulating or suppressing activity, hematopoietic regulating
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC activin/inhibin activity, chemostatic/chemokinetic activity,
CC cachectin/tumour invasion suppressor activity, tissue growth activity,
CC tumour inhibition activity or other activities
SQ Sequence 1097 BP; 365 A; 209 C; 202 G; 321 T;
Query Match 100.0%; Score 9; DB 47; Length 1097;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 382 ttctcagaa 390
|||||
Cp 33 ttctcagaa 25
RESULT 180
ID Q96139 standard; cDNA; 1100 BP.
AC Q96139;
DT 16-JAN-1996 (first entry)
DE Prolectin cDNA.
KW Prolectin; cDNA; vaccine; augment; bacterins; attenuated vaccine;
KW live vaccine; virus; immune response; ss.
OS Homo sapiens.
PN W09521625-A1.
PD 17-AUG-1995.
PF 14-FEB-1995; U01866.
PR 14-FEB-1994; US-196350.
PA (GENZ) GENZYME CORP.
PI Kaplan J, Mosicki R, Richards S;
PI WPI: 95-292943/38.
DR Use of prolectin or prolectin cDNA - for enhancing the immune
PT response of an animal to an infectious disease vaccine
PS Claim 8; Page 14-15; 22pp; English.
CC A composition comprising prolectin or prolectin cDNA can be used for
CC enhancing the immune response of an animal to an infectious disease
CC vaccine. The composition can be used to enhance the effectiveness
CC of vaccines which are considered "weak" e.g. Bacterins and
CC attenuated live or killed virus products.
SQ Sequence 1100 BP; 324 A; 282 C; 202 G; 292 T;
Query Match 100.0%; Score 9; DB 16; Length 1100;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 977 ttctcagaa 985
|||||
Oy 25 ttctcagaa 33
RESULT 181
ID Q72807 standard; cDNA; 1110 BP.
AC Q72807;
DT 22-JUN-1995 (first entry)
DE DNA encoding stem region of A2/Aichi/2/68 influenza virus.
KW Conserved peptide; stem region; hemagglutinin; HA; H1N1; H2N2;
KW subtype; human; influenza A virus; immunogenic artificial peptide;
KW antigen; vaccine; infection; ss.
OS Human influenza A virus.
FH Key Location/Qualifiers
FT cds 34; 1077
FT /*tag- a
PN EP-621339-A.
PD 26-OCT-1994.
PF 20-APR-1994; 302819.
PR 20-APR-1993; JP-115216.
PR 16-MAR-1994; JP-070194.
PA (TAKI) TAKARA SHUZO CO LTD.
PI Isegawa Y, Okuno Y, Sasao F, Ueda S;

DR WPI: 94-325949/41.
P-PSDB: R63591.
PT Human influenza A virus haemagglutinin polypeptide(s) - useful in
PT Influenza A vaccine composition
PS Claim 17; Page 59; 68pp; English.
CC This sequence encodes the stem region of the haemagglutinin (HA) molecule
CC of the A2/Aichi/2/68 strain of human influenza A virus. This antigenic
CC molecule contains the conserved peptides derived from the stem region
CC of the H1N1 and H2N2 subtypes of human influenza A virus. Immunogenic
CC polypeptides such as this are antigenically equivalent to the stem region
CC of the HA molecule of influenza A virus. This artificial peptide may be
CC used as a vaccine for prophylaxis of influenza A virus infection.
SQ Sequence 1110 BP; 374 A; 216 C; 258 G; 262 T;
Query Match 100.0%; Score 9; DB 12; Length 1110;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 619 ttctcagaa 627
|||||
Cp 33 ttctcagaa 25
RESULT 182
ID Q21023 standard; DNA; 1112 BP.
AC Q21023;
DT 29-MAY-1992 (first entry)
DE mg1b-delta-alpha1M fusion gene.
DE galactose binding protein; microglobulin; immunosay; antibody;
KW tubular kidney function; ss.
FH Key Location/Qualifiers
FT cds 1..1101
FT /*tag- a
FT /label- fusion_protein
PN EP-470565-A.
PD 12-FEB-1992.
PF 05-AUG-1991; 113161.
PR 06-AUG-1990; DE-024919.
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
PI Kopeckí E, Klein C, Mangold D, Stock W, Schlupfenbacher R;
PI WPI: 92-050631/07.
DR P-PSDB: R21085.
PT Peptide with alpha-1-micro:globulin binding characteristics -
PT used in conjunction with specific antibodies for determining
PT alpha-1 micro:globulin in body fluids
PS Example 2.2.2; Page 9 + 20; 26pp; German.
CC Oligonucleotides 1-6 (Q21017-22) are used in the prodn. of
CC alpha1M(aa54-87)-alpha1M(aa161-183). The ligated sequences form
CC a 200 bp gene, which is inserted in the HindIII/BamHI pBR322 vector
CC fragment (pBR322-delta-alpha1M). The pBR322-delta-alpha1M vector
CC insert is then inserted into pVBI/EH to form pVBI/EH-delta-alpha1M.
CC Example 2.2 describes the prod. of galactose binding protein(aa1-306)
CC -alpha1M(aa54-87)-alpha1M(aa161-183) fusion protein.
CC The protein and antibodies to the protein are used for immunosay
CC of human microglobulin (for assessing tubular kidney function).
CC See also Q21015-23.
SQ Sequence 1112 BP; 310 A; 268 C; 313 G; 221 T;
Query Match 100.0%; Score 9; DB 3; Length 1112;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 ttctcagaa 128
|||||
Cp 33 ttctcagaa 25
RESULT 183
ID V30002 standard; DNA; 1119 BP.
AC V30002;
DT 13-AUG-1998 (first entry)
DE DNA encoding a Staphylococcus aureus sugar kinase.
KW Sugar Kinase; Staphylococcus aureus WCUH 29 strain NCIMB 40771;

[illegible]

Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 936 ttctcagaa 944
|||||

Cp 33 ttctcagaa 25

RESULT 187

ID V55785 standard: cDNA; 1151 BP.

AC V55785;

DT 23-MAR-1999 (first entry)

DE Mouse B7-2 antigen coding sequence.

KM B7-2 antigen; mammalian tumour cell; T cell costimulation; CD28 ligand;

KW CTLA4 ligand; therapy; T-cell response; mouse; ss.

OS Mus sp.

FH Location/Qualifiers

FT CDS 99..1028

FT CDS /*tag- a

US5861310-A.

PD 19-JAN-1999.

PF 30-MAY-1995; 456104.

PR 03-NOV-1993; US-147773.

PA (DAND) DANA FARBER CANCER INST INC.

PI Freeman GJ, Gray GS, Nadler LM;

DR WPI: 99-130394/11.

DR P-PSDB: W73639.

PT Tumour cell transfected to express B7-2 molecule - useful for tumour

PT therapy by stimulating T-cell response

PS Disclosure; Column 31-34; 27pp; English.

CC This sequence encodes the mouse B7-2 antigen, which can be used in the

CC method of the invention. The method is for transfecting an isolated

CC mammalian tumour cell with an exogenous nucleic acid molecule encoding a

CC mammalian B7-2 molecule, where the B7-2 molecule is expressed in the

CC tumour cell, is capable of costimulating a T cell and is capable of

CC binding a CD28 or CTLA4 ligand. The method is useful for treating tumours

CC by stimulating a T-cell response against tumour cells in vivo.

SQ Sequence 1151 BP; 353 A; 258 C; 255 G; 285 T;

Query Match 100.0%; Score 9; DB 60; Length 1151;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

Db 1091 ttctcagaa 1099
|||||

Cp 33 ttctcagaa 25

RESULT 188

ID Q81366 standard: cDNA; 1163 BP.

AC Q81366;

DT 21-AUG-1995 (first entry)

DE Murine B lymphocyte antigen B7-2.

KW B lymphocyte antigen; mb7-2; ss.

OS Mus musculus.

FH Location/Qualifiers

FT CDS 111..1040

FT CDS /*tag- a

WO9503408-A.

PD 02-FEB-1995.

PF 26-JUL-1994; U08433.

PR 26-JUL-1993; US-101624.

PR 19-AUG-1993; US-109393.

PR 03-NOV-1993; US-147773.

PA (DAND) DANA FARBER CANCER INST INC.

PI Freeman GJ, Gray GS, Greenfield E, Nadler LM;

DR WPI: 95-075236/10.

DR P-PSDB: R67988.

PT Nucleic acids encoding CTLA4/CD28 counter receptor, B7-2 - useful
for enhancing or suppressing T-cell mediated immune responses

PS Claim 7: Page 106-107; 175pp; English.

CC A cDNA library was constructed in the pCDM8 vector using poly A+

CC RNA from activated M12 cells (a murine B cell tumour line).

CC COS cells were transfected with the M12 murine B cell library

CC DNA and screened for expression of B7-2. 6 clones were strongly

CC positive. The B7-2 cDNA insert in clone 4 was sequenced in the

CC pCDM8 expression vector employing the following strategy: initial

CC sequencing was performed using sequencing primers T7 (Q81352) and

CC CDMB8 (Q81353). DNA sequence obtd. using these primers was used to

CC design additional sequencing primers (see Q81367-Q81370). The cycle

CC of sequencing and selection of additional primers was continued

CC until the murine B7-2 cDNA was completely sequenced on both

CC strands. A murine B7-2 clone (mb7-2, clone 4) was obtd.

CC (see Q81366/R67988). The predicted protein exhibits many features

CC common to other type I Ig superfamily membrane proteins. Cleavage

CC of the signal peptide would result in a murine B7-2 membrane bound

CC protein having an unmodified mol. wt. of approx. 32 kDa. The

CC extracellular domain contains nine potential N-linked glycosylation

CC sites. Glycosylation may increase the mol. wt. to about 50-70 kDa.

CC E. coli transfected with a vector contg. a cDNA insert encoding

CC mb7-2 clone 4 was deposited with the ATCC on August 18th 1993 as

CC Accession No. 69388.

SQ Sequence 1163 BP; 354 A; 265 C; 258 G; 286 T;

Query Match 100.0%; Score 9; DB 14; Length 1163;
Best Local Similarity 100.0%; Pred. No. 1.37e+03;

Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

Db 1103 ttctcagaa 1111
|||||

Cp 33 ttctcagaa 25

RESULT 189

ID V37401 standard: DNA; 1167 BP.

AC V37401;

DT 13-OCT-1998 (first entry)

DE Streptococcus pneumoniae coding region.

KW coding region; ORF; open reading frame; antibacterial;

KM infection; prevention; meningitis; ss.

OS Streptococcus pneumoniae.

FH Location/Qualifiers

FT CDS 79..492

FT CDS /*tag- a

FT CDS /*product- unknown

WO9819689-A1.

PD 14-MAY-1998.

PF 27-OCT-1997; U19226.

PR 01-NOV-1996; US-029930.

PA (SMK) SMITHKLINE BEECHAM CORP.

PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,

DR Reid RH, Zarits PN;

DR WPI: 98-286586/25.

DR P-PSDB: W61013.

PT New isolated nucleic acids from Streptococcus pneumoniae - useful,

PT e.g. for identifying anti-bacterial(s) for treatment and prevention

PT of meningitis

PS Claim 1: Page 107-108; 130pp; English.

CC The sequence is that of a coding region isolated from

CC S. pneumoniae. Its encoded protein, or agonists of it,

CC may be useful as an antibacterial for treatment or

CC prevention of infection, specifically caused by S. pneumoniae

CC (particularly meningitis) but possibly also Helicobacter

CC pylori (ulcers and gastric cancer). It may be of particular

CC use before insertion of an in-dwelling device or any other

CC invasive procedure. The protein, or nucleic acid encoding

CC it, can also be used in vaccines to induce a cellular

CC and/or humoral immune response, or to screen for other

CC antibacterials. The DNA may also contain flanking sequences

CC that are potential sources of control elements for bacterial

CC gene expression. Detecting a sequence encoding the protein

CC can be used diagnostically, e.g. to detect a mutation for

CC serotyping or classifying infectious agents.

SQ Sequence 1167 BP; 385 A; 206 C; 247 G; 327 T;

Query Match 100.0%; Score 9; DB 46; Length 1167;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 833 tctcagaa 841
 |||||||
 25 tctcagaa 33

RESULT 190
 ID Q9950 standard; DNA; 1176 BP.
 AC Q9950;
 DT 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A encoding DNA.
 KW 11-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 KW anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection; ss.
 OS Homo sapiens.
 FH Key
 FT Location/Qualifiers
 FT 108..1175
 FT CDS
 FT /tag= a
 FT /product= IL-8A_receptor
 FT W09525126-A1.
 FT 21-SEP-1995.
 FT 09-MAR-1995: U03032.
 FT 15-MAR-1994: US-210250.
 FT 02-MAY-1994: US-237937.
 FT (REPR) REPLIGEN CORP.
 FT (UYO-) UNIV BOSTON.
 FT Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Witt DF.
 DR WPI: 95-336945/43.
 DR P-PSDB: R80951.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 PT useful for treating inflammatory disorders, for detecting
 PT neutrophil(s) and for isolating IL-8 receptor from 11q.mixt.
 PS Disclosure: Fig 2A-B; 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q995049-52). The A subtype receptor (IL-8A) is
 CC a high affinity receptor and the B subtype receptor (IL-8B) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 1176 BP; 247 A; 333 C; 279 G; 317 T;

Query Match 100.0%; Score 9; DB 17; Length 1176;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1108 tctcagaa 1116
 |||||||
 33 tctcagaa 25

RESULT 191
 ID V84163 standard; DNA; 1180 BP.
 AC V84163;
 DT 30-MAR-1999 (first entry)
 DE Streptococcus uberis plasminogen activator protein Paua DNA.
 KW Paua; plasminogen activator; streptokinase; mastitis; vaccine;
 KW ss.
 OS Streptococcus uberis.
 FH Key
 FT Location/Qualifiers
 FT CDS
 FT 120..980
 FT /tag= a
 FT /note= "this region is specifically claimed in claim 5"
 FT 120..194
 FT sig_peptide
 FT /tag= b

FT mat_peptide 195..977
 FT /tag= c
 FT /note= "this region is specifically claimed in claim 3"
 FT terminator 978..1088
 FT /tag= d

Db 556 tctcagaa 564
 |||||||
 33 tctcagaa 25

RESULT 192
 ID Q85933 standard; cDNA; 1181 BP.
 AC Q85933;
 DT 02-NOV-1995 (first entry)
 DE cDNA encoding Dac gV (clone 259).
 KW Lollum perenne; Lol PV; Dactylis glomerata; Dac gV; epitope;
 KW sensitivity; ryegrass pollen allergen; ss.
 OS Dactylis glomerata.
 FH Key
 FT Location/Qualifiers
 FT misc-feature 1..25
 FT /tag= a
 FT /note= "adaptor sequence"
 FT cds 53..964
 FT /tag= b
 FT /product= immature Dac gV protein
 FT signal_peptide 1..124
 FT /tag= c
 FT mat_peptide 125..961
 FT /tag= d

W09506728-A.
 PN 09-MAR-1995.
 PF 05-AUG-1994: U09024.
 PF 13-AUG-1993: US-106016.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Griffith J, Kuo M, Lugman M;

WPI: 95-115444/15.
 DR P-SDS: R71507.
 CC Lactum perenne Lol pv and Dactylis glomerata Dac gv epitope(s)
 PT and DNA - for treating sensitivity to rye-grass pollen allergen
 PT or an immunologically cross-reactive allergen.
 PS Example 5, Page 73-74; 110pp; English.
 CC This is the cDNA sequence of clone 259, encoding Dac gv derived from a
 CC lambda gt10 library. The nucleotide sequence from 700 to 1181 is only
 CC preliminary and some bases may be misidentified (sic). Dac gv, isolated
 CC from Dactylis glomerata, is a major protein allergen related to ryegrass
 CC pollen allergens, eg. Lol pv. Peptides comprising at least one T cell
 CC epitope derived from the Lol pv protein are claimed, and can be used to
 CC treat or diagnose sensitivity to ryegrass pollen in an individual or to
 CC pollen proteins that are immunologically related to Lol pv, such as
 CC Dac gv.
 CC Sequence 1181 BP; 260 A; 420 C; 316 G; 184 T;
 Query Match 100.0%; Score 9; DB 15; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1132 ttctgaga 1140
 |||||||
 Qy 25 ttctgaga 33
 RESULT 193
 ID V84164 standard; DNA; 1181 BP.
 AC V84164;
 DT 30-MAR-1999 (first entry)
 DE Streptococcus uberis plasminogen activator protein Paua DNA.
 KW Paua; plasminogen activator; streptokinase; mastitis; vaccine;
 KW ss.
 OS Streptococcus uberis
 FH Key Location/Qualifiers
 FT CDS 121..981
 FT /*tag- a
 FT /*note- "this region is specifically claimed in claim 5"
 FT 121..195
 FT /*tag- b
 FT 196..978
 FT mat_peptide
 FT /*tag- C
 FT /*note- "this region is specifically claimed in claim 3"
 FT 979..1089
 FT /*tag- d
 FT terminator
 FT EP-887410-A2.
 PD 30-DEC-1998;
 PF 17-JUN-1998; 304759.
 PR 23-JUN-1997; US-050565.
 PR (PR12) PEIZER INC.
 PI Leigh JA, Rosey EL, Yancey RJ;
 DR WPI: 99-047877/05.
 DR P-SDS: W88240.
 PT New Streptococcus uberis plasminogen activator (Paua)
 PT polynucleotides and polypeptides - useful as diagnostic probes and
 PT as vaccines to protect against mastitis caused by bacterial
 PT infection
 PS Claim 6; Page 30-31; 50pp; English.
 CC This claimed polynucleotide, the coding region of which is also
 CC claimed, encodes a new Streptococcus uberis plasminogen activator,
 CC termed Paua (see W88239), that is involved in the development of
 CC mastitis in mammals. The paua gene was amplified from S. uberis
 CC strain 95-140 chromosomal DNA using probes (see V84166-68) based on
 CC isolated peptides of Paua and by screening of transformant libraries
 CC for clearing activity on skim milk plates. A paua gene was also
 CC isolated from S. uberis strain c216 (see V84163). The paua
 CC polynucleotides are useful for recombinant production of Paua
 CC polypeptides. They are also useful for designing diagnostic primers
 CC and probes for detection of bacterial disease, including mastitis-
 CC causing streptococci, particularly S. uberis, by amplification and
 CC detection of paua polynucleotides (claimed). The Paua polypeptides
 CC are useful as vaccines to protect mammals against mastitis
 CC (claimed) and other diseases and conditions by stimulating the

CC production of anti-plasminogen activator antibodies, as pathogens
 CC use plasminogen activator to convert inactive plasmin in bovine
 CC milk to active plasmin, for use in growth. This is preferable to
 CC prior art methods using chemical antibiotics
 CC Sequence 1181 BP; 415 A; 198 C; 193 G; 375 T;
 Query Match 100.0%; Score 9; DB 60; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 557 ttctcaga 565
 |||||||
 Cp 33 ttctcaga 25
 RESULT 194
 ID T62556 standard; DNA; 1185 BP.
 AC T62556;
 DT 21-MAY-1997 (first entry)
 DE Human IP-10 gene.
 KW Human IP-10 gene.
 KW IP-10; chemokine; inhibitor; endothelial cell; endothelioma;
 KW tumour; Kaposi's sarcoma; sclerosing endothelial sarcoma;
 KW hemangioma; neoplasm; arteriosclerosis; inflammation; fibrosis;
 KW keloid; proliferative retinal disorder; traumatic injury;
 KW ischaemia; heparin sulphate proteoglycan receptor;
 KW fibroblast growth factor; transforming growth factor-beta;
 KW gene therapy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 80..376
 FT /*tag- a
 FT 1168..1173
 FT poly_a_signal
 FT /*tag- b
 FT WO9700691-A1.
 PD 09-JAN-1997.
 PF 21-JUN-1996; U10726.
 PR 22-JUN-1995; US-493638.
 PR (HARD) HARVARD COLLEGE.
 PI Leder P, Luster A;
 DR WPI: 97-087169/08.
 DR P-SDS: W14307.
 PT Modulation of endothelial cell growth - using IP-10 or IP-10
 PT inhibitors, used for treating e.g. tumours, inflammation or
 PT traumatic or ischemic injury
 PS Disclosure; Page 52; 70pp; English.
 CC A DNA molecule (T62556) codes for human IP-10 (W14307), a 10 kDa
 CC secreted protein that binds to the heparin sulphate proteoglycan
 CC receptor and which is capable of inhibiting endothelial cell
 CC proliferation. IP-10 nucleic acids can be used to prepare
 CC recombinant IP-10, esp. as a non-aggregating fusion protein with
 CC alkaline phosphatase as fusion partner. IP-10, or a nucleic acid
 CC encoding it (provided e.g. by a viral vector), can be used to
 CC inhibit endothelial cell proliferation, e.g. in the treatment of
 CC endotheliomas or arteriosclerosis, and to inhibit fibroblast
 CC growth factor or transforming growth factor-beta activity.
 CC Sequence 1185 BP; 389 A; 234 C; 210 G; 352 T;
 Query Match 100.0%; Score 9; DB 29; Length 1185;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 731 ttctcaga 739
 |||||||
 Cp 33 ttctcaga 25
 RESULT 195
 ID V59981 standard; CDNA to mRNA; 1189 BP.
 AC V59981;
 DT 25-NOV-1998 (first entry)
 DE Nucleic acid Tci encoding human homologue of rat acyl-CoA synthetase.
 KW Human homologue; yeast Rhd50; Drosophila Septin-2; Acyl-CoA synthetase;
 KW Immunomodulatory activity; identification; activated T-cell; ds.

OS Homo sapiens.
 PN M09838306-A1.
 PD 03-SEP-1998.
 PF 27-FEB-1997; U03159.
 PR 27-FEB-1997; WO-003159.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Dolganov G;
 DR WPI: 98-481207/41.
 PT Novel human immunomodulatory poly:peptide(s) - have homology to the
 PT yeast RAD50 or Drosophila Septin-2 proteins
 PS Disclosure: Page 63; 155pp; English.
 CC V595981-89 represent multiple isoforms of a human homologue of the rat
 CC brain long chain acyl-CoA synthetase. The present sequence has 90%
 CC overall homology to the rat gene, and is expressed in activated
 CC T-cells, testis, and foetal liver tissues. The specification
 CC also describes sequences encoding human homologues of the
 CC yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
 CC proteins have immunomodulatory activity. The nucleic acids and
 CC proteins can be used to identify activated T-cells in a sample
 CC population. They can also be used to isolate and identify sequences
 CC encoding other proteins or other compounds having immunomodulatory
 CC activity.
 SQ Sequence 1189 BP; 307 A; 281 C; 235 G; 366 T;
 Query Match 100.0%; Score 9; DB 49; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 1173 ttctcagaa 1181
 |||||
 Cp 33 ttctcagaa 25

RESULT 196
 ID 073730 standard; cDNA; 1197 BP.
 AC 073730;
 DT 26-APR-1995 (first entry)
 DE Rape abscission/dehiscence protein PSAC40 cDNA.
 KW Rape; PSAC40: abscission; dehiscence; transgenic plant;
 KM crop improvement; ss.
 OS Brassica napus.
 FH Key Location/Qualifiers
 FT cds 92..1052
 FT /*tag= a
 PN M09423043-A.
 PD 13-OCT-1994.
 PF 31-MAR-1994; G00689.
 PR 31-MAR-1993; GB-006726.
 PA (NICK-) NICKERSON BIOCEM LTD.
 PI Coupe SA, Isaac PG, Roberts JA;
 DR WPI: 94-333203/41.
 DR P-PSDB; R60800.
 PT Recombinant and isolated nucleic acids - encode enzymes and
 PT proteins involved in plant abscission or dehiscence.
 PS Disclosure: Fig 8; 39pp; English.
 CC The cDNA encodes a rape (cv. Rafal) protein involved in abscission
 CC or dehiscence, and can be used to manipulate plant abscission and
 CC reduce/prevent pod dehiscence, especially to regulate abscission of
 CC pollen from anthers, to generate male sterile plants and for hybrid
 CC seed production.
 SQ Sequence 1197 BP; 389 A; 259 C; 218 G; 331 T;
 Query Match 100.0%; Score 9; DB 12; Length 1197;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 679 ttctcagaa 687
 |||||
 Oy 25 ttctcagaa 33

RESULT 197
 ID V37033 standard; DNA; 1204 BP.

AC V37033;
 DT 04-SEP-1998 (first entry)
 DE Streptococcus pneumoniae sequence used to design primers and probes.
 KW Detection; bacterial antibiotic resistance gene; bacteria;
 KM fungal species; identification; ds.
 OS Streptococcus pneumoniae.
 PN M09820157-A2.
 PD 14-MAY-1998.
 PR 04-NOV-1997; CA0829.
 PR 04-NOV-1996; US-743637.
 PA (IDI-) IDI INFECTIO DIAGNOSTIC INC.
 PI Bergeron MG, Ouellette M, Picard FT, Roy PH;
 DR WPI: 98-286967/25.
 PT Use of oligo:nucleotide primers and probes - for detection,
 PT identification and quantification of bacteria, fungi and bacterial
 PT antibiotic resistance gene(s)
 PS Claim 10; Pages 85-86; 167pp; English.
 CC The present sequence was used to design primers and probes which
 CC are used in the course of the invention. The specification describes the
 CC use of probes and/or amplification primers which are specific, ubiquitous
 CC and sensitive for determining the presence and amount of nucleic acids
 CC from a bacterial antibiotic resistance gene and specific bacterial and
 CC fungal species in any sample suspected of containing the bacterial or
 CC fungal nucleic acids, where each of the nucleic acid or variant or part
 CC comprises a selected target region hybridisable with the probes or
 CC primers. The method of use comprises contacting the sample with the
 CC probes or primers and detecting the presence of hybridised probes or
 CC amplified products as an indication of the presence of the specific
 CC bacterial or fungal species and bacterial antibiotic resistance genes.
 CC The methods and products can be used to detect and identify the bacterial
 CC and fungal species and genera and determine the bacterial resistance to
 CC antibiotics
 SQ Sequence 1204 BP; 390 A; 202 C; 302 G; 310 T;
 Query Match 100.0%; Score 9; DB 44; Length 1204;
 Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 954 ttctcagaa 962
 |||||
 Oy 25 ttctcagaa 33

RESULT 198
 ID M40167 standard; DNA; 1212 BP.
 AC M40167;
 DT 08-JAN-1992 (first entry)
 DE Sequence encoding galactosidase fusion protein containing Plasmodium
 DE knowlesi CS protein immunogenic region.
 KW Malaria; vaccine; epitope; ss.
 OS Plasmodium knowlesi.
 FH Key Location/Qualifiers
 FT cds 1..1212
 FT /*tag= a
 FT misc_feature 631..666
 FT /*tag= b
 FT /*note= "claimed SQ"
 PN M08402922-A.
 PD 02-AUG-1984.
 PF 27-JAN-1984; U00144.
 PR 28-JAN-1983; GB-002349.
 PR 27-JAN-1984; GB-002211.
 PR 27-JAN-1984; GB-002186.
 PA (UNY-) NEW YORK UNIV.
 PI Colman DR, Ellis J, Godson GN, Nussenzweig RS, Nussenzweig,
 PI VN, Schlesinger DH, Svec PS, Zavala F;
 DR P-PSDB; P40213.
 DR Peptide epitope of sporozoite CS protein of plasmodium - useful
 PT for vaccine to confer protection against malaria infections
 PS Disclosure: Fig 1; 67pp; English.
 CC The inventors claim a peptide comprising an epitope of a sporozoite
 CC CS protein of a member of the genus Plasmodium (see P40213) and a

CC DNA encoding it (see M40167). The peptides may be used in vaccines
 CC to confer protection against malaria infections. The protective CS
 CC sporozoite antigens of Plasmodium possess a single immunodominant
 CC region composed of repetitions of the same epitope.
 SQ Sequence 1212 BP; 351 A; 287 C; 337 G; 237 T;

Query Match 100.0%; Score 9; DB 3; Length 1212;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 ttctcagaa 293

Cp 33 ttctcagaa 25

RESULT 199
 ID V18169 standard; DNA; 1218 BP.
 AC V18169;
 DT 07-DEC-1998 (first entry)
 DE Gulf War veteran subject #1 PCR clone #7A.
 KW Picornavirus; poliovirus; enterovirus; vaccine; diagnosis;
 KW chronic illness; polio; Gulf War syndrome; multiple myeloma;
 KW prostate cancer; autoimmune disease; neuromuscular disease;
 KW Parkinson's disease; ss.
 OS Enterovirus.
 FH Key Location/Qualifiers
 FT primer_bind Complement (1..18)
 FT /*tag= a
 FT /note= "primer PG02"
 FT 1199..1218
 FT primer_bind /*tag= b
 FT /note= "primer PG03"

PN W09814617-A1.
 PD 09-APR-1998.
 PF 03-OCT-1997; U17880.
 PR 04-OCT-1996; US-026762.
 PA (CHRO-) CHRONIX BIOMEDICAL.
 PI Urooviltz HB;
 DR WPI: 98-240111/21.
 PT Method for screening for target human nucleic acids - useful for,
 PT e.g. diagnosing and monitoring disease such as cancer or multiple
 PT myeloma
 PS Example 1; Page 38; 56pp; English.
 CC Clone #7A is a PCR product obtained from amplification of nucleic
 CC acids from serum samples of a Gulf War veteran (subject #1),
 CC diagnosed with Gulf War syndrome, using primers PG02 and PG03 (see
 CC V18154-55) that are based on poliovirus 5'-nontranslated and p2
 CC regions. #7A is longer than the expected PCR product and shows no
 CC sequence identity to any known sequence (see also W18164-74). The
 CC invention provides hybridisation assays for the detection of target
 CC human nucleic acids in a biological sample. The methods are
 CC particularly useful for the early diagnosis of chronic illnesses
 CC such as cancer, multiple myeloma, autoimmune diseases, and
 CC neurodegenerative diseases (all claimed). The invention is also
 CC directed to the detection of non-poliovirus nucleic acids (e.g.,
 CC in vaccine polio preparations) and recombinants between polio and
 CC other viruses. Such recombinants have been detected in Gulf War
 CC syndrome patients and in patients diagnosed with multiple sclerosis
 CC and prostate cancer.
 SQ Sequence 1218 BP; 378 A; 287 C; 249 G; 302 T;

Query Match 100.0%; Score 9; DB 49; Length 1218;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1169 ttctcagaa 1177

Cp 33 ttctcagaa 25

RESULT 200
 ID T42978 standard; cDNA; 1221 BP.
 AC T42978;

DT 12-FEB-1997 (first entry)
 DE Human pancreas-derived serpin cDNA.
 KW Pancreas-derived serpin; PDS; serine protease inhibitor;
 KW inflammation; antiinflammatory; diagnosis; gene therapy; ss.
 OS Homo sapiens.
 PN W09634957-A1.
 PD 07-NOV-1996.
 PF 01-MAY-1996; U06137.
 PR 02-MAY-1995; US-434881.
 PA (INCY-) INCYTE PHARM INC.
 PI Braxton SM, Diep D, Wilde CG;
 DR WPI: 96-518311/51.
 DR P-PDSB; W06202.

PT DNA encoding pancreas derived serpin - useful in diagnosis and
 PT treatment of pancreatic inflammation and disease
 PS Claim 2; Page 24; 36pp; English.
 CC A cDNA clone (T42978), designated incyte 222689, codes for human
 CC pancreas-derived serpin (PDS) (W06202), a novel serine protease
 CC inhibitor. It was identified in a human pancreas library via
 CC phagemid cDNA clone isolation, sequencing and homology searching
 CC against a DNA sequence database. The cDNA can be used in the
 CC prodn. of recombinant PDS in host cells, and to design probes or
 CC primers useful in diagnostic tests for conditions associated with
 CC pancreatic inflammation. Antisense sequences can be used to treat
 CC pancreatic inflammation or disease.
 SQ Sequence 1221 BP; 417 A; 212 C; 249 G; 343 T;

Query Match 100.0%; Score 9; DB 25; Length 1221;

Best Local Similarity 100.0%; Pred. No. 1.37e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 659 ttctcagaa 667

Cp 25 ttctcagaa 33

Search completed: Thu Oct 28 00:33:22 1999
 Job time : 198 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Oct 28 00:33:44 1999; Maspar time 3.73 Seconds

Tabular output not generated. 208.910 Million cell updates/sec

Title: >US-08-963-288-1

Description: (25-33) from US08963288.seq

Perfect Score: 9

N.A. Sequence: 25 ttcttgagaa 33

Comp: aagactctt

Scoring table: TABLE default

Gap 6

Mismatch SMD : Dbase 0; Query 0

Searched: 165359 seqs, 43243793 bases x 2

Post-processing: Minimum Match 100%

Listing first 200 summaries

Database: n-issued
1:5A_COMB 2:5B_COMB 3:5C_COMB 4:PCT9_COMB 5:backfiles1

Statistics: Mean 4.019; Variance 1.621; scale 2.479

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	9	100.0	9	4	PCT-US95-0 Sequence 35, Applicati	3.96e+02
2	9	100.0	9	4	PCT-US95-1 Sequence 34, Applicati	3.96e+02
3	9	100.0	9	3	US-08-410- Sequence 35, Applicati	3.96e+02
4	9	100.0	9	2	US-08-369- Sequence 34, Applicati	3.96e+02
5	9	100.0	9	2	US-08-411- Sequence 13, Applicati	3.96e+02
6	9	100.0	9	3	US-08-410- Sequence 101, Applicat	3.96e+02
7	9	100.0	19	4	PCT-US95-0 Sequence 57, Applicati	3.96e+02
8	9	100.0	19	2	US-08-411- Sequence 101, Applicati	3.96e+02
9	9	100.0	19	4	PCT-US95-0 Sequence 100, Applicat	3.96e+02
10	9	100.0	19	3	US-08-410- Sequence 100, Applicat	3.96e+02
11	9	100.0	19	2	US-08-411- Sequence 56, Applicati	3.96e+02
12	9	100.0	20	3	US-08-723- Sequence 21, Applicati	3.96e+02
13	9	100.0	20	4	PCT-US94-1 Sequence 20, Applicati	3.96e+02
14	9	100.0	22	3	US-08-564- Sequence 18, Applicati	3.96e+02
15	9	100.0	22	3	US-08-537- Sequence 18, Applicati	3.96e+02
16	9	100.0	22	1	US-08-198- Sequence 18, Applicati	3.96e+02
17	9	100.0	26	4	PCT-US93-0 Sequence 141, Applicat	3.96e+02
18	9	100.0	26	1	US-08-379- Sequence 663, Applicat	3.96e+02
19	9	100.0	26	4	PCT-US93-0 Sequence 234, Applicat	3.96e+02
20	9	100.0	26	1	US-08-379- Sequence 667, Applicat	3.96e+02

21	9	100.0	26	1	US-08-379- Sequence 553, Applicat	3.96e+02
22	9	100.0	26	4	PCT-US93-0 Sequence 181, Applicat	3.96e+02
23	9	100.0	26	4	PCT-US93-0 Sequence 176, Applicat	3.96e+02
24	9	100.0	26	1	US-08-379- Sequence 558, Applicat	3.96e+02
25	9	100.0	26	4	PCT-US93-0 Sequence 226, Applicat	3.96e+02
26	9	100.0	27	3	US-08-431- Sequence 5, Applicatio	3.96e+02
27	9	100.0	28	4	PCT-US95-1 Sequence 38, Applicatio	3.96e+02
28	9	100.0	28	4	PCT-US95-1 Sequence 37, Applicati	3.96e+02
29	9	100.0	28	3	US-08-431- Sequence 23, Applicati	3.96e+02
30	9	100.0	29	3	US-08-431- Sequence 3, Applicatio	3.96e+02
31	9	100.0	31	2	US-08-117- Sequence 51, Applicati	3.96e+02
32	9	100.0	32	4	PCT-US96-0 Sequence 11, Applicati	3.96e+02
33	9	100.0	32	3	US-08-425- Sequence 20, Applicati	3.96e+02
34	9	100.0	40	1	US-07-938- Sequence 1, Applicatio	3.96e+02
35	9	100.0	40	1	US-07-938- Sequence 3, Applicatio	3.96e+02
36	9	100.0	40	1	US-07-938- Sequence 30, Applicati	3.96e+02
37	9	100.0	41	4	PCT-US92-0 Sequence 18, Applicati	3.96e+02
38	9	100.0	41	4	PCT-US92-0 Sequence 18, Applicati	3.96e+02
39	9	100.0	42	1	US-08-564- Sequence 17, Applicati	3.96e+02
40	9	100.0	42	3	US-08-537- Sequence 17, Applicati	3.96e+02
41	9	100.0	42	1	US-08-198- Sequence 17, Applicati	3.96e+02
42	9	100.0	42	1	US-08-431- Sequence 1, Applicatio	3.96e+02
43	9	100.0	54	3	US-08-758- Sequence 558, Applicat	3.96e+02
44	9	100.0	54	1	US-08-390- Sequence 1036, Applicat	3.96e+02
45	9	100.0	54	2	US-08-435- Sequence 6, Applicatio	3.96e+02
46	9	100.0	54	2	US-08-532- Sequence 6, Applicatio	3.96e+02
47	9	100.0	60	1	US-07-938- Sequence 9, Applicatio	3.96e+02
48	9	100.0	60	1	PCT-US96-0 Sequence 15, Applicatio	3.96e+02
49	9	100.0	63	4	US-07-982- Sequence 2, Applicatio	3.96e+02
50	9	100.0	63	3	US-07-982- Sequence 16, Applicati	3.96e+02
51	9	100.0	65	3	US-07-982- Sequence 9, Applicatio	3.96e+02
52	9	100.0	66	3	US-07-982- Sequence 9, Applicatio	3.96e+02
53	9	100.0	74	5	5200336-2 Patent No. 5200336.	3.96e+02
54	9	100.0	77	1	US-08-477- Sequence 5, Applicatio	3.96e+02
55	9	100.0	77	1	US-08-477- Sequence 5, Applicatio	3.96e+02
56	9	100.0	77	1	US-08-477- Sequence 5, Applicatio	3.96e+02
57	9	100.0	80	2	US-08-441- Sequence 18, Applicati	3.96e+02
58	9	100.0	80	2	US-07-938- Sequence 10, Applicati	3.96e+02
59	9	100.0	80	1	US-07-938- Sequence 11, Applicati	3.96e+02
60	9	100.0	130	1	US-07-938- Sequence 4, Applicatio	3.96e+02
61	9	100.0	130	1	US-07-938- Sequence 6, Applicatio	3.96e+02
62	9	100.0	132	1	US-08-343- Sequence 12, Applicati	3.96e+02
63	9	100.0	156	3	US-07-700- Sequence 28, Applicati	3.96e+02
64	9	100.0	161	1	US-07-858- Sequence 16, Applicati	3.96e+02
65	9	100.0	171	1	US-08-398- Sequence 43, Applicati	3.96e+02
66	9	100.0	171	1	US-08-398- Sequence 20, Applicati	3.96e+02
67	9	100.0	171	1	US-07-858- Sequence 22, Applicati	3.96e+02
68	9	100.0	171	1	US-07-858- Sequence 21, Applicati	3.96e+02
69	9	100.0	192	3	US-08-151- Sequence 21, Applicati	3.96e+02
70	9	100.0	192	2	US-08-060- Sequence 37, Applicati	3.96e+02
71	9	100.0	192	2	US-08-153- Sequence 21, Applicati	3.96e+02
72	9	100.0	197	1	US-07-858- Sequence 17, Applicati	3.96e+02
73	9	100.0	197	1	US-07-858- Sequence 18, Applicati	3.96e+02
74	9	100.0	261	4	PCT-US95-0 Sequence 29, Applicati	3.96e+02
75	9	100.0	267	3	US-08-928- Sequence 5, Applicatio	3.96e+02
76	9	100.0	297	2	US-08-406- Sequence 5, Applicatio	3.96e+02
77	9	100.0	331	1	US-08-171- Sequence 1, Applicatio	3.96e+02
78	9	100.0	342	3	US-08-917- Sequence 1, Applicatio	3.96e+02
79	9	100.0	363	3	US-08-470- Sequence 6, Applicatio	3.96e+02
80	9	100.0	495	4	US-08-470- Sequence 6, Applicatio	3.96e+02
81	9	100.0	521	4	PCT-US92-1 Sequence 13, Applicati	3.96e+02
82	9	100.0	521	1	US-07-998- Sequence 4, Applicatio	3.96e+02
83	9	100.0	651	1	US-08-171- Sequence 16, Applicati	3.96e+02
84	9	100.0	666	4	PCT-US94-0 Sequence 3, Applicatio	3.96e+02
85	9	100.0	666	4	PCT-US94-0 Sequence 1, Applicatio	3.96e+02
86	9	100.0	670	4	PCT-US94-0 Sequence 7, Applicatio	3.96e+02

C 94	9	100.0	688	3	US-08-139-	Sequence 9, Applicatio	3.96e+02
C 95	9	100.0	688	4	PCT-US93-1	Sequence 9, Applicatio	3.96e+02
C 96	9	100.0	756	4	PCT-US96-0	Sequence 2, Applicatio	3.96e+02
C 97	9	100.0	756	1	US-08-216-	Sequence 1, Applicatio	3.96e+02
C 98	9	100.0	790	2	US-08-117-	Sequence 7, Applicatio	3.96e+02
C 99	9	100.0	792	2	US-08-565-	Sequence 5, Applicatio	3.96e+02
C 100	9	100.0	792	2	US-08-565-	Sequence 5, Applicatio	3.96e+02
C 101	9	100.0	796	3	US-08-967-	Sequence 119, Applicat	3.96e+02
C 102	9	100.0	796	4	PCT-US92-0	Sequence 2, Applicatio	3.96e+02
C 103	9	100.0	796	4	US-08-375-	Sequence 2, Applicatio	3.96e+02
C 104	9	100.0	819	2	US-08-407-	Sequence 1, Applicatio	3.96e+02
C 105	9	100.0	819	2	US-08-967-	Sequence 118, Applicat	3.96e+02
C 106	9	100.0	850	2	US-08-485-	Sequence 4, Applicatio	3.96e+02
C 107	9	100.0	850	2	US-08-247-	Sequence 11, Applicatio	3.96e+02
C 108	9	100.0	858	2	US-08-346-	Sequence 15, Applicatio	3.96e+02
C 109	9	100.0	858	4	PCT-US91-0	Sequence 15, Applicatio	3.96e+02
C 110	9	100.0	861	3	US-08-406-	Sequence 1, Applicatio	3.96e+02
C 111	9	100.0	875	5	US-08-650-	Sequence 1, Applicatio	3.96e+02
C 112	9	100.0	875	5	5258283-1	Patent No. 5258283.	3.96e+02
C 113	9	100.0	879	1	US-08-158-	Sequence 3, Applicatio	3.96e+02
C 114	9	100.0	892	3	US-08-801-	Sequence 4, Applicatio	3.96e+02
C 115	9	100.0	918	4	PCT-US94-0	Sequence 9, Applicatio	3.96e+02
C 116	9	100.0	918	4	PCT-US94-0	Sequence 58, Applicatio	3.96e+02
C 117	9	100.0	928	4	PCT-US96-0	Sequence 5, Applicatio	3.96e+02
C 118	9	100.0	958	3	US-08-632-	Sequence 9, Applicatio	3.96e+02
C 119	9	100.0	958	3	US-08-888-	Sequence 12, Applicatio	3.96e+02
C 120	9	100.0	969	2	US-08-310-	Sequence 12, Applicatio	3.96e+02
C 121	9	100.0	972	1	US-08-243-	Sequence 2, Applicatio	3.96e+02
C 122	9	100.0	987	1	US-08-352-	Sequence 57, Applicatio	3.96e+02
C 123	9	100.0	1000	1	US-07-965-	Sequence 3, Applicatio	3.96e+02
C 124	9	100.0	1021	1	US-08-459-	Sequence 17, Applicatio	3.96e+02
C 125	9	100.0	1040	1	US-08-247-	Sequence 1, Applicatio	3.96e+02
C 126	9	100.0	1047	2	US-08-843-	Sequence 5, Applicatio	3.96e+02
C 127	9	100.0	1060	1	US-07-940-	Sequence 11, Applicatio	3.96e+02
C 128	9	100.0	1090	4	PCT-US96-0	Sequence 4, Applicatio	3.96e+02
C 129	9	100.0	1093	3	US-08-352-	Sequence 95, Applicatio	3.96e+02
C 130	9	100.0	1100	1	US-08-196-	Sequence 2, Applicatio	3.96e+02
C 131	9	100.0	1106	1	US-08-041-	Sequence 4, Applicatio	3.96e+02
C 132	9	100.0	1106	1	US-08-041-	Sequence 6, Applicatio	3.96e+02
C 133	9	100.0	1106	1	US-08-041-	Sequence 8, Applicatio	3.96e+02
C 134	9	100.0	1110	1	US-08-630-	Sequence 57, Applicatio	3.96e+02
C 135	9	100.0	1110	1	US-08-229-	Sequence 57, Applicatio	3.96e+02
C 136	9	100.0	1129	1	PCT-US94-1	Sequence 1, Applicatio	3.96e+02
C 137	9	100.0	1129	2	US-08-470-	Sequence 1, Applicatio	3.96e+02
C 138	9	100.0	1146	1	US-08-487-	Sequence 1, Applicatio	3.96e+02
C 139	9	100.0	1151	3	US-08-456-	Sequence 3, Applicatio	3.96e+02
C 140	9	100.0	1151	3	PCT-US95-0	Sequence 20, Applicatio	3.96e+02
C 141	9	100.0	1176	4	PCT-US95-0	Sequence 2, Applicatio	3.96e+02
C 142	9	100.0	1185	3	US-08-493-	Sequence 5, Applicatio	3.96e+02
C 143	9	100.0	1188	4	PCT-US95-0	Sequence 6, Applicatio	3.96e+02
C 144	9	100.0	1188	3	US-08-921-	Sequence 6, Applicatio	3.96e+02
C 145	9	100.0	1189	3	US-08-592-	Sequence 56, Applicatio	3.96e+02
C 146	9	100.0	1196	3	US-08-864-	Sequence 2, Applicatio	3.96e+02
C 147	9	100.0	1221	1	US-08-434-	Sequence 1, Applicatio	3.96e+02
C 148	9	100.0	1226	1	US-08-334-	Sequence 2, Applicatio	3.96e+02
C 149	9	100.0	1226	1	US-08-090-	Sequence 26, Applicatio	3.96e+02
C 150	9	100.0	1226	1	US-08-398-	Sequence 26, Applicatio	3.96e+02
C 151	9	100.0	1227	3	US-09-031-	Sequence 18, Applicatio	3.96e+02
C 152	9	100.0	1227	3	US-08-847-	Sequence 18, Applicatio	3.96e+02
C 153	9	100.0	1227	3	US-09-031-	Sequence 17, Applicatio	3.96e+02
C 154	9	100.0	1227	3	US-08-847-	Sequence 17, Applicatio	3.96e+02
C 155	9	100.0	1228	3	US-09-031-	Sequence 16, Applicatio	3.96e+02
C 156	9	100.0	1228	3	US-08-847-	Sequence 16, Applicatio	3.96e+02
C 157	9	100.0	1228	3	US-09-031-	Sequence 14, Applicatio	3.96e+02
C 158	9	100.0	1228	3	US-08-847-	Sequence 14, Applicatio	3.96e+02
C 159	9	100.0	1242	4	PCT-US93-0	Sequence 4, Applicatio	3.96e+02
C 160	9	100.0	1261	4	PCT-US95-0	Sequence 12, Applicatio	3.96e+02
C 161	9	100.0	1273	1	US-07-794-	Sequence 1, Applicatio	3.96e+02
C 162	9	100.0	1273	1	US-07-794-	Sequence 2, Applicatio	3.96e+02
C 163	9	100.0	1284	4	PCT-US96-0	Sequence 28, Applicatio	3.96e+02
C 164	9	100.0	1290	4	US-08-906-	Sequence 1, Applicatio	3.96e+02
C 165	9	100.0	1296	4	PCT-US94-1	Sequence 3, Applicatio	3.96e+02
C 166	9	100.0	1328	1	US-07-679-	Sequence 14, Applicatio	3.96e+02

167	9	100.0	1338	1	US-07-651-	Sequence 1, Applicatio	3.96e+02
168	9	100.0	1340	1	US-07-679-	Sequence 16, Applicatio	3.96e+02
169	9	100.0	1341	1	US-08-185-	Sequence 26, Applicatio	3.96e+02
C 170	9	100.0	1341	1	PCT-US94-0	Sequence 26, Applicatio	3.96e+02
C 171	9	100.0	1341	1	US-08-180-	Sequence 26, Applicatio	3.96e+02
C 172	9	100.0	1353	3	US-08-625-	Sequence 3, Applicatio	3.96e+02
C 173	9	100.0	1375	3	US-08-736-	Sequence 1, Applicatio	3.96e+02
C 174	9	100.0	1407	4	PCT-US96-0	Sequence 3, Applicatio	3.96e+02
C 175	9	100.0	1453	1	US-08-252-	Sequence 1, Applicatio	3.96e+02
C 176	9	100.0	1461	4	PCT-US96-0	Sequence 27, Applicatio	3.96e+02
C 177	9	100.0	1482	1	US-07-768-	Sequence 1, Applicatio	3.96e+02
C 178	9	100.0	1501	4	PCT-US95-1	Sequence 5, Applicatio	3.96e+02
C 179	9	100.0	1501	4	US-08-349-	Sequence 1, Applicatio	3.96e+02
C 180	9	100.0	1510	4	PCT-US92-0	Sequence 1, Applicatio	3.96e+02
C 181	9	100.0	1512	1	US-07-768-	Sequence 3, Applicatio	3.96e+02
C 182	9	100.0	1550	3	US-08-466-	Sequence 17, Applicatio	3.96e+02
C 183	9	100.0	1550	3	US-08-475-	Sequence 17, Applicatio	3.96e+02
C 184	9	100.0	1600	1	US-08-252-	Sequence 5, Applicatio	3.96e+02
C 185	9	100.0	1610	1	US-07-768-	Sequence 1, Applicatio	3.96e+02
C 186	9	100.0	1614	3	US-08-399-	Sequence 2, Applicatio	3.96e+02
C 187	9	100.0	1635	3	US-08-709-	Sequence 1, Applicatio	3.96e+02
C 188	9	100.0	1634	3	US-08-460-	Sequence 1, Applicatio	3.96e+02
C 189	9	100.0	1635	5	5198359-9	Patent No. 5198359.	3.96e+02
C 190	9	100.0	1656	5	5449756-9	Patent No. 5449756.	3.96e+02
C 191	9	100.0	1656	5	5198359-5	Patent No. 5198359.	3.96e+02
C 192	9	100.0	1673	3	US-08-476-	Sequence 4, Applicatio	3.96e+02
C 193	9	100.0	1673	3	US-08-833-	Sequence 4, Applicatio	3.96e+02
C 194	9	100.0	1673	3	US-08-306-	Sequence 4, Applicatio	3.96e+02
C 195	9	100.0	1673	4	PCT-US91-0	Sequence 4, Applicatio	3.96e+02
C 196	9	100.0	1701	3	US-08-884-	Sequence 40, Applicatio	3.96e+02
C 197	9	100.0	1722	1	US-08-247-	Sequence 5, Applicatio	3.96e+02
C 198	9	100.0	1732	2	US-08-417-	Sequence 17, Applicatio	3.96e+02
C 199	9	100.0	1749	5	5258283-8	Patent No. 5258283.	3.96e+02
C 200	9	100.0	1757	3	US-08-453-	Sequence 14, Applicatio	3.96e+02

ALIGNMENTS

RESULT 1
PCT-US95-04477-35 STANDARD; DNA; UNC; 9 BP.
AC xxxxxx

DE Sequence 35, Application PC/TUS9504477.
CC Sequence 35, Application PC/TUS9504477
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS RESPONSIVE TO
CC TITLE OF INVENTION: CYTOKINES AND METHODS FOR THEIR USE
CC NUMBER OF SEQUENCES: 165
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/04477
CC FILING DATE:
CC PRIORITY APPLICATION DATA:
CC CLASSIFICATION:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994
CC INFORMATION FOR SEQ ID NO: 35:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 9 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
CC SEQUENCE 9 BP; 3 A; 2 C; 1 G; 3 T; 0 OTHER.
SO

Query Match 100.0%; Score 9; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTCTCAGAA 9
|||||
Cp 33 ttctcagaa 25

RESULT 2
ID PCT-US95-17025-34 STANDARD; DNA; UNC; 9 BP.
AC xxxxxx

DE Sequence 34, Application PC/TUS9517025
CC Sequence 34, Application PC/TUS9517025
CC GENERAL INFORMATION:
CC APPLICANT: James E. Darnell, Jr.
CC APPLICANT: Zilong Wen
CC APPLICANT: Curt M. Horvath
CC APPLICANT: Zhong Zhong
CC TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
CC TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
CC NUMBER OF SEQUENCES: 39
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Klauber & Jackson
CC STREET: 411 Hackensack Avenue
CC CITY: Hackensack
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07601

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/17025
CC FILING DATE: 28-DEC-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/369,796
CC FILING DATE: 06-JAN-1995

CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jackson Esq., David A.
CC REGISTRATION NUMBER: 26,742
CC REFERENCE/DOCKET NUMBER: 600-1-116
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201 487-5800
CC TELEFAX: 201 343-1684
CC TELEX: 133521

CC INFORMATION FOR SEQ ID NO: 34:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 9 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA synthetic probe
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO

CC SEQUENCE 9 BP: 3 A; 2 C; 1 G; 3 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTCTCAGAA 9
|||||
Cp 33 ttctcagaa 25

RESULT 3
ID US-08-410-779B-35 STANDARD; DNA; UNC; 9 BP.
AC xxxxxx

DE Sequence 35, Application US/08410779B
CC Sequence 35, Application US/08410779B
CC Patent No. 5814517
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS
CC TITLE OF INVENTION: RESPONSIVE TO CYTOKINES AND METHODS FOR THEIR USE
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LIGAND PHARMACEUTICALS INCORPORATED
CC STREET: 9393 TOWNE CENTRE DRIVE
CC CITY: SAN DIEGO
CC STATE: CALIFORNIA
CC COUNTRY: US
CC ZIP: 92121

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/410,779B
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994

CC ATTORNEY/AGENT INFORMATION:
CC NAME: JURGENSEN, THOMAS E
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0013A, US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 35:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 9 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
CC SEQUENCE 9 BP: 3 A; 2 C; 1 G; 3 T; 0 OTHER.
SQ

Query Match 100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTCTCAGAA 9
|||||
Cp 33 ttctcagaa 25

RESULT 4
ID US-08-369-796-34 STANDARD; DNA; UNC; 9 BP.
AC xxxxxx

DE Sequence 34, Application US/08369796
CC Sequence 34, Application US/08369796
CC Patent No. 5716622
CC GENERAL INFORMATION:
CC APPLICANT: James E. Darnell, Jr.
CC APPLICANT: Zilong Wen
CC APPLICANT: Curt M. Horvath
CC APPLICANT: Zhong Zhong
CC TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
CC TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEIN
CC NUMBER OF SEQUENCES: 39
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Klauber & Jackson

CC SEQUENCE 9 BP: 3 A; 2 C; 1 G; 3 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC STREET: 411 Hackensack Avenue
CC CITY: Hackensack
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07601
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/369,796
CC FILING DATE: 06-JAN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jackson Esq., David A.
CC REGISTRATION NUMBER: 26,742
CC REFERENCE/DOCKET NUMBER: 600-1-116
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201 487-5800
CC TELEFAX: 201 343-1684
CC TELEX: 133521
CC INFORMATION FOR SEQ ID NO: 34:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 9 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA synthetic probe
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 9 BP: 3 A; 2 C; 1 G; 3 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 TTCTCAGAA 9
| | | | |
Cp 33 ttctcagaa 25
RESULT 5
ID US-08-411-020-13 STANDARD; DNA: UNC; 9 BP.
AC xxxxxx
DT Sequence 13, Application US/08411020
DE Sequence 13, Application US/08411020
CC Patent No. 5712094
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC APPLICANT: CHAN, SHIN-SHAY TIAN
CC TITLE OF INVENTION: METHODS AND ASSOCIATED REAGENTS FOR
CC TITLE OF INVENTION: DETECTING MODULATORS OF CYTOKINE ACTION
CC NUMBER OF SEQUENCES: 59
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ligand Pharmaceuticals Incorporated
CC STREET: 9393 Towne Centre Drive
CC CITY: San Diego
CC STATE: California
CC COUNTRY: US
CC ZIP: 92121
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/411,020
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Jurgensen, Thomas E.
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0030.US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 9 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
CC SEQUENCE 9 BP: 3 A; 2 C; 1 G; 3 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 TTCTCAGAA 9
| | | | |
Cp 33 ttctcagaa 25
RESULT 6
ID US-08-410-779B-101 STANDARD; DNA: UNC; 19 BP.
AC xxxxxx
DT Sequence 101, Application US/08410779B
DE Sequence 101, Application US/08410779B
CC Patent No. 5814517
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS
CC TITLE OF INVENTION: RESPONSIVE TO CYTOKINES AND METHODS FOR THEIR USE
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LIGAND PHARMACEUTICALS INCORPORATED
CC STREET: 9393 TOWNE CENTRE DRIVE
CC CITY: SAN DIEGO
CC STATE: CALIFORNIA
CC COUNTRY: US
CC ZIP: 92121
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/410,779B
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: JURGENSEN, THOMAS E.
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0013A.US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 101:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,

CC DESCRIPTION: SYNTHETIC DNA"
SQ SEQUENCE 19 BP; 6 A; 4 C; 5 G; 4 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTGAGAA 16
|||||
25 ttctgagaa 33

RESULT 7
ID PCT-US95-04477-101 STANDARD; DNA; UNC; 19 BP.
AC xxxxxx

DE Sequence 101, Application PC/TUS9504477
CC Sequence 101, Application PC/TUS9504477
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS RESPONSIVE TO
CC NUMBER OF SEQUENCES: 165
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/04477
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994
CC INFORMATION FOR SEQ ID NO: 101:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
SQ SEQUENCE 19 BP; 6 A; 4 C; 5 G; 4 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTGAGAA 16
|||||
25 ttctgagaa 33

RESULT 8
ID US-08-411-020-57 STANDARD; DNA; UNC; 19 BP.
AC xxxxxx

DE Sequence 57, Application US/08411020
CC Sequence 57, Application US/08411020
CC Patent No. 5712094
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC APPLICANT: CHAN, SHIN-SHAY TIAN
CC TITLE OF INVENTION: METHODS AND ASSOCIATED REAGENTS FOR
CC NUMBER OF SEQUENCES: 59
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ligand Pharmaceuticals Incorporated
CC STREET: 9393 Towne Centre Drive
CC CITY: San Diego

CC STATE: California
CC COUNTRY: US
CC ZIP: 92121

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/411,020
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jurgensen, Thomas E.
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0030-US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 57:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
SQ SEQUENCE 19 BP; 6 A; 4 C; 5 G; 4 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTGAGAA 16
|||||
25 ttctgagaa 33

RESULT 9
ID PCT-US95-04477-100 STANDARD; DNA; UNC; 19 BP.
AC xxxxxx

DE Sequence 100, Application PC/TUS9504477
CC Sequence 100, Application PC/TUS9504477
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS RESPONSIVE TO
CC NUMBER OF SEQUENCES: 165
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/04477
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
SQ SEQUENCE 19 BP; 4 A; 5 C; 6 G; 6 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTCAGAA 16
|||||
CP 33 ttctcagaa 25

RESULT 10
ID US-08-410-779B-100 STANDARD; DNA; UNC; 19 BP.
AC xxxxxx
DE Sequence 100, Application US/08410779B
CC Sequence 100, Application US/08410779B
CC Patent No. 5814517
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC TITLE OF INVENTION: DNA SPACER REGULATORY ELEMENTS
CC TITLE OF INVENTION: RESPONSIVE TO CYTOKINES AND METHODS FOR THEIR USE
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LIGAND PHARMACEUTICALS INCORPORATED
CC STREET: 9393 TOWNE CENTRE DRIVE
CC CITY: SAN DIEGO
CC STATE: CALIFORNIA
CC COUNTRY: US
CC ZIP: 92121
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/410.779B
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/228,935
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: JURGENSEN, THOMAS E.
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0013A.US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
SO SEQUENCE 19 BP; 4 A; 5 C; 4 G; 6 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTCAGAA 16
|||||
CP 33 ttctcagaa 25

RESULT 11
ID US-08-411-020-56 STANDARD; DNA; UNC; 19 BP.
AC xxxxxx
DE Sequence 56, Application US/08411020

CC Sequence 56, Application US/08411020
CC Patent No. 5712094
CC GENERAL INFORMATION:
CC APPLICANT: SEIDEL, H. MARTI
CC APPLICANT: LAMB, I. PETER
CC APPLICANT: CHAN, SHIN-SHAY TIAN
CC TITLE OF INVENTION: METHODS AND ASSOCIATED REAGENTS FOR
CC TITLE OF INVENTION: DETECTING MODULATORS OF CYTOKINE ACTION
CC NUMBER OF SEQUENCES: 59
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ligand Pharmaceuticals Incorporated
CC STREET: 9393 Towne Centre Drive
CC CITY: San Diego
CC STATE: California
CC COUNTRY: US
CC ZIP: 92121
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/411,020
CC FILING DATE: 27-MAR-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jurgensen, Thomas E.
CC REGISTRATION NUMBER: 34,195
CC REFERENCE/DOCKET NUMBER: 016-0030.US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 550-7675
CC TELEFAX: (619) 535-3906
CC INFORMATION FOR SEQ ID NO: 56:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 19 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "OTHER NUCLEIC ACID,
CC DESCRIPTION: SYNTHETIC DNA"
SO SEQUENCE 19 BP; 4 A; 5 C; 4 G; 6 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 TTCTCAGAA 16
|||||
CP 33 ttctcagaa 25

RESULT 12
ID US-08-723-415B-21 STANDARD; DNA; UNC; 20 BP.
AC xxxxxx
DE Sequence 21, Application US/08723415B
CC Sequence 21, Application US/08723415B
CC Patent No. 5859199
CC GENERAL INFORMATION:
CC APPLICANT: Lathangue, Nicholas B.
CC APPLICANT: delaluna, Susana
CC TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
CC TITLE OF INVENTION: THEREOF
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: NIXON & VANDERHYE P.C.
CC STREET: 1100 NO. 5659199th Gléde Rd. 8th floor
CC CITY: Arlington
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22201-4741
CC COMPUTER READABLE FORM:

1
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/723,415B
CC FILING DATE: 30-SEP-1996
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: GB 9610195.1
CC FILING DATE: 15-MAY-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Crawford, Arthur R.
CC REGISTRATION NUMBER: 25,327
CC REFERENCE/DOCKET NUMBER: 117-220
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-816-4000
CC TELEFAX: 703-816-4100
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 20 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "PRIMER"
CC SEQUENCE 20 BP: 7 A; 3 C; 4 G; 6 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 TTCTGAGAA 9
OY 25 ttctgagaa 33
RESULT 13
ID PCT-US94-10261A-20 STANDARD: DNA; UNC: 20 BP.
AC xxxxxx
DE Sequence 20, Application PC/TUS9410261A
CC Sequence 20, Application PC/TUS9410261A
CC GENERAL INFORMATION:
CC APPLICANT: Carter, William G.
CC APPLICANT: Gili, Susanna A.
CC APPLICANT: Ryan, Maureen C.
CC TITLE OF INVENTION: Epiligrin, an Epithelial Ligand for
CC TITLE OF INVENTION: Integrins
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness
CC STREET: 1420 Fifth Avenue
CC CITY: Seattle
CC STATE: WA
CC COUNTRY: USA
CC ZIP: 98101-8100
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/10261A
CC FILING DATE: 02-SEP-1994
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Shelton, Dennis K.
CC REGISTRATION NUMBER: 26,997
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 682-8100
CC TELEFAX: (206) 224-0779
CC INFORMATION FOR SEQ ID NO: 20:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 20 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA to mRNA
CC DESCRIPTION: PCR primer; see TABLE 1
CC HYPOTHETICAL: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC SEQUENCE 20 BP: 9 A; 3 C; 4 G; 4 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 4 TTCTGAGAA 12
OY 25 ttctgagaa 33
RESULT 14
ID US-08-564-955-18 STANDARD: DNA; UNC: 22 BP.
AC xxxxxx
DE Sequence 18, Application US/08564955
CC Sequence 18, Application US/08564955
CC Patent No. 5811238
CC GENERAL INFORMATION:
CC APPLICANT: STEMMER, WILLEM P.C.
CC APPLICANT: CRAMER, ANDREAS M.
CC TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
CC TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AN
CC NUMBER OF SEQUENCES: 67
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
CC STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: U.S.A.
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/564,955
CC FILING DATE: 30-NOV-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/198,431
CC FILING DATE: 17-FEB-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/537,874
CC FILING DATE: 30-OCT-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02126
CC FILING DATE: 17-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: DUNN, TRACY J.
CC REGISTRATION NUMBER: 34,587
CC REFERENCE/DOCKET NUMBER: 165285-014611US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 326-2400
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 22 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

SEQ SEQUENCE 22 BP; 8 A; 4 C; 3 G; 7 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 22;

Best Local Similarity 100.0%; Pred. No. 3.96e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 TTCTGAGAA 21

0Y 25 ttctgagaa 33

RESULT 15
ID US-08-537-874-18 STANDARD; DNA; UNC; 22 BP.

AC xxxxxx

DE Sequence 18, Application US/08537874

CC Sequence 18, Application US/08537874

CC Patent No. 5830721

CC GENERAL INFORMATION:

CC APPLICANT: Stemmer, Willem P.C.

CC TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation

CC TITLE OF INVENTION: and Reassembly

CC NUMBER OF SEQUENCES: 62

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Townsend and Townsend and Crew LLP

CC STREET: Two Embarcadero Center, 8th Floor

CC CITY: San Francisco

CC STATE: CA

CC COUNTRY: USA

CC ZIP: 94111

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/537,874

CC FILING DATE:

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: EP PCT/US95/02126

CC FILING DATE: 17-FEB-1995

CC APPLICATION NUMBER: US 08/198,431

CC FILING DATE: 17-FEB-1994

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Liedeschuetz, Joe

CC REGISTRATION NUMBER: 37,505

CC REFERENCE/DOCKET NUMBER: 018097-014610

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 415-576-0200

CC TELEFAX: 415-576-0300

CC INFORMATION FOR SEQ ID NO: 18:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 22 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC SEQUENCE 22 BP; 8 A; 4 C; 3 G; 7 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 22;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 TTCTGAGAA 21

0Y 25 ttctgagaa 33

RESULT 16

ID US-08-198-431-18 STANDARD; DNA; UNC; 22 BP.

AC xxxxxx

DT

DE Sequence 18, Application US/08198431

CC Sequence 18, Application US/08198431

CC Patent No. 5605793

CC GENERAL INFORMATION:

CC APPLICANT: Stemmer, Willem P.C.

CC TITLE OF INVENTION: Methods for In Vitro Recombination

CC NUMBER OF SEQUENCES: 49

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Burns, Doane, Swecker & Mathis

CC STREET: 699 Prince Street

CC CITY: Alexandria

CC STATE: Virginia

CC COUNTRY: USA

CC ZIP: 22313

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/198,431

CC FILING DATE: 17-FEB-1994

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Mool, Leslie

CC REGISTRATION NUMBER: 37,047

CC REFERENCE/DOCKET NUMBER: 000324-003

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 415-854-7400

CC TELEFAX: 415-854-8275

CC INFORMATION FOR SEQ ID NO: 18:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 22 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (oligonucleotide)

CC SEQUENCE 22 BP; 8 A; 4 C; 3 G; 7 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 22;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 TTCTGAGAA 21

0Y 25 ttctgagaa 33

RESULT 17

ID PCT-US93-00977-141 STANDARD; DNA; UNC; 26 BP.

AC xxxxxx

DT

DE Sequence 141, Application PC/TUS9300977

CC Sequence 141, Application PC/TUS9300977

CC GENERAL INFORMATION:

CC TITLE OF INVENTION: METHOD AND REAGENT FOR MEASURING MESSENGER RNA

CC NUMBER OF SEQUENCES: 711

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Knobbe, Martens, Olson, and Bear

CC STREET: 620 Newport Center Dr. Sixteenth Floor

CC CITY: Newport Beach

CC STATE: CA

CC COUNTRY: USA

CC ZIP: 92660

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US93/00977

CC FILING DATE: 19930129

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC - REFERENCE/DOCKET NUMBER: HITACHI.006H
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 141:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 6 A; 6 C; 4 G; 10 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 TTCTGAGAA 22
|||||
0Y 25 ttctcgagaa 33

RESULT 18
ID US-08-379-078-663 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx

DE Sequence 663, Application US/08379078
CC Sequence 663, Application US/08379078
CC Patent No. 5639612
CC GENERAL INFORMATION:
CC APPLICANT: Mitsuhashi, Masato
CC APPLICANT: Cooper, Allan
CC TITLE OF INVENTION: Gene Detection System
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
CC STREET: 620 Newport Center Drive 16th Floor
CC CITY: Newport Beach
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible.
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/379,078
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/974,406
CC FILING DATE: 12-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC REFERENCE/DOCKET NUMBER: HITACHI.011CP2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 663:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO

CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 10 A; 4 C; 6 G; 6 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTGAGAA 13
|||||
Cp 33 ttctcgagaa 25

RESULT 19
ID PCT-US93-00977-234 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx

DE Sequence 234, Application PC/TUS9300977
CC Sequence 234, Application PC/TUS9300977
CC GENERAL INFORMATION:
CC TITLE OF INVENTION: METHOD AND REAGENT FOR MEASURING MESSENGER RNA
CC NUMBER OF SEQUENCES: 711
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Knobbe, Martens, Olson, and Bear
CC STREET: 620 Newport Center Dr. Sixteenth Floor
CC CITY: Newport Beach
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/00977
CC FILING DATE: 19930129
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC REFERENCE/DOCKET NUMBER: HITACHI.006H
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 234:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 10 A; 4 C; 6 G; 6 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTGAGAA 13
|||||
Cp 33 ttctcgagaa 25

RESULT 20
ID US-08-379-078-667 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx

DE Sequence 667, Application US/08379078
CC Sequence 667, Application US/08379078
CC Patent No. 5639612
CC GENERAL INFORMATION:
CC APPLICANT: Mitsuhashi, Masato

CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 10 A; 4 C; 6 G; 6 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 5 TTCACGAA 13
|||||
CP 33 ttctcagaa 25
RESULT 23
ID PCT-US93-00977-176 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx
DE Sequence 176, Application PC/TUS9300977
CC Sequence 176, Application PC/TUS9300977
CC GENERAL INFORMATION: METHOD AND REAGENT FOR MEASURING MESSENGER RNA
CC TITLE OF INVENTION: 711
CC NUMBER OF SEQUENCES: 711
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Knobb, Martens, Olson, and Bear
CC STREET: 620 Newport Center Dr. Sixteenth Floor
CC CITY: Newport Beach
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/00977
CC FILING DATE: 19930129
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC REFERENCE/DOCKET NUMBER: HITACHI.006H
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 176:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 10 A; 4 C; 6 G; 6 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 5 TTCACGAA 13
|||||
CP 33 ttctcagaa 25
RESULT 24
ID US-08-379-078-558 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx
DE Sequence 558, Application US/08379078
CC Sequence 558, Application US/08379078
CC Patent No. 5639612
CC GENERAL INFORMATION:

CC APPLICANT: Mitsuhashi, Masato
CC APPLICANT: Cooper, Allan
CC TITLE OF INVENTION: Gene Detection System
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: KNOBB, MARTENS, OLSON AND BEAR
CC STREET: 620 Newport Center Drive 16th Floor
CC CITY: Newport Beach
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/379,078
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/974,406
CC FILING DATE: 12-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC REFERENCE/DOCKET NUMBER: HITACHI.011CP2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 558:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 26 BP; 10 A; 4 C; 6 G; 6 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 5 TTCACGAA 13
|||||
CP 33 ttctcagaa 25
RESULT 25
ID PCT-US93-00977-226 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx
DE Sequence 226, Application PC/TUS9300977
CC Sequence 226, Application PC/TUS9300977
CC GENERAL INFORMATION:
CC TITLE OF INVENTION: METHOD AND REAGENT FOR MEASURING MESSENGER RNA
CC NUMBER OF SEQUENCES: 711
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Knobb, Martens, Olson, and Bear
CC STREET: 620 Newport Center Dr. Sixteenth Floor
CC CITY: Newport Beach
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 92660
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/00977

CC FILING DATE: 19930129
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Altman, Daniel E.
CC REGISTRATION NUMBER: 34,115
CC REFERENCE/DOCKET NUMBER: HITACHI.006H
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 714-760-0404
CC TELEFAX: 714-760-9502
CC INFORMATION FOR SEQ ID NO: 226:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 26 BP: 10 A; 4 C; 6 G; 6 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 5 TTCTCAGAA 13
|||||||
Cp 33 ttctcagaa 25
RESULT 26
ID US-08-431-527A-5 STANDARD; DNA; UNC; 27 BP.
AC xxxxxx
DE Sequence 5, Application US/08431527A
CC Sequence 5, Application US/08431527A
CC Patent No. 5843650
CC GENERAL INFORMATION:
CC APPLICANT: David Segev
CC TITLE OF INVENTION: NO. 5843650-enzymatic method for detecting nucleic acid seq
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
CC STREET: 2940 Birchtree space lane
CC CITY: Silver Spring
CC STATE: Maryland
CC COUNTRY: United States of America
CC ZIP: 20906
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
CC COMPUTER: twinhead slimnote-890TX
CC OPERATING SYSTEM: MS DOS version 6.2,
CC SOFTWARE: Word for Windows version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/431.527A
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Friedmam, Mark M.
CC REGISTRATION NUMBER: 33,883
CC REFERENCE/DOCKET NUMBER: 128/8
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 972-3-6938541
CC TELEFAX: 972-3-6938542
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 27
CC TYPE: nucleic acid
CC STRANDEDNESS: single

CC TOPOLOGY: linear
SO SEQUENCE 27 BP: 5 A; 9 C; 5 G; 7 T; 1 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 16 TTCTCAGAA 24
|||||||
Cp 33 ttctcagaa 25
RESULT 27
ID PCT-US95-15716-38 STANDARD; DNA; UNC; 28 BP.
AC xxxxxx
DE Sequence 38, Application PC/TUS9515716
CC Sequence 38, Application PC/TUS9515716
CC GENERAL INFORMATION:
CC APPLICANT: Berdoz, Jose
CC APPLICANT: Kraehenbuhl, Jean Pierre
CC TITLE OF INVENTION: PCR AMPLIFICATION OF REARRANGED GENOMIC
CC TITLE OF INVENTION: VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
CC NUMBER OF SEQUENCES: 108
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street, Suite 3100
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30B
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/15716
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/348,548
CC FILING DATE: 01-DEC-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 06132/009001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-5070
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 38:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 28 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA
CC SEQUENCE 28 BP: 7 A; 7 C; 6 G; 8 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 18 TTCTCAGAA 26
|||||||
Cp 33 ttctcagaa 25
RESULT 28
ID PCT-US95-15716-37 STANDARD; DNA; UNC; 28 BP.
AC xxxxxx
DE Sequence 37, Application PC/TUS9515716

Sequence 37, Application PC/TUS9515716
GENERAL INFORMATION:
APPLICANT: Berdoz, Jose
-APPLICANT: Kraehenbuhl, Jean Pierre
TITLE OF INVENTION: PCR AMPLIFICATION OF REARRANGED GENOMIC
TITLE OF INVENTION: VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
NUMBER OF SEQUENCES: 108
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15716
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,548
FILING DATE: 01-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06132/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-5070
TELEX: 200154
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE 28 BP: 8 A; 6 C; 6 G; 8 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 18 TTCTCAGAA 26
|||||
CP 33 ttctcagaa 25
RESULT 29
ID PCT-US95-15716-23 STANDARD: DNA; UNC; 28 BP.
AC xxxxxx
DE Sequence 23, Application PC/TUS9515716
CC Sequence 23, Application PC/TUS9515716
CC GENERAL INFORMATION:
CC APPLICANT: Berdoz, Jose
CC APPLICANT: Kraehenbuhl, Jean Pierre
CC TITLE OF INVENTION: PCR AMPLIFICATION OF REARRANGED GENOMIC
CC TITLE OF INVENTION: VARIABLE REGIONS OF IMMUNOGLOBULIN GENES
CC NUMBER OF SEQUENCES: 108
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street, Suite 3100
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15716
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348,548
FILING DATE: 01-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06132/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-5070
TELEX: 200154
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE 28 BP: 8 A; 8 C; 6 G; 6 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 28;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 18 TTCTCAGAA 26
|||||
CP 33 ttctcagaa 25
RESULT 30
ID US-08-431-527A-3 STANDARD: DNA; UNC; 29 BP.
AC xxxxxx
DE Sequence 3, Application US/08431527A
CC Sequence 3, Application US/08431527A
CC Patent No. 5843650
CC GENERAL INFORMATION:
CC APPLICANT: David Segev
CC TITLE OF INVENTION: No. 5843650-enzymatic method for detecting nucleic acid
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
CC STREET: 2940 Birchtree space lane
CC CITY: Silver Spring
CC STATE: Maryland
CC COUNTRY: United States of America
CC ZIP: 20906
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
CC COMPUTER: Twinhead Slimnote-890TX
CC OPERATING SYSTEM: MS DOS version 6.2,
CC OPERATING SYSTEM: Windows version 3.11
CC SOFTWARE: Word for Windows version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/431,527A
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Friedman, Mark M.
CC REGISTRATION NUMBER: 33,883
CC REFERENCE/DOCKET NUMBER: 128/8
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 972-3-6938541
CC TELEFAX: 972-3-6938542
CC TELEX:

CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 29
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 29 BP; 9 A; 5 C; 9 G; 5 T; 1 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 4 TTCTGAGAA 12
| | | | | | | | | |
Oy 25 ttctgagaa 33
RESULT 31
ID US-08-117-083-51 STANDARD; DNA; UNC; 31 BP.
AC xxxxxx
DE Sequence 51, Application US/08117083
CC Sequence 51, Application US/08117083
CC Patent No. 5719054
CC GENERAL INFORMATION:
CC APPLICANT: Boursnell, Michael E.
CC APPLICANT: Inglis, Stephen C.
CC APPLICANT: Munro, Alan J.
CC TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
CC NUMBER OF SEQUENCES: 70
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Walter H. Dreyer
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/117,083
CC FILING DATE: 10-SEP-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dreyer, Walter H.
CC REGISTRATION NUMBER: 24,190
CC REFERENCE/DOCKET NUMBER: A-58783
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-781-1989
CC TELEFAX: 415-398-3249
CC TELEX: 910 277299
CC INFORMATION FOR SEQ ID NO: 51:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 31 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC SEQUENCE 31 BP; 7 A; 6 C; 9 G; 9 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 17 TTCTGAGAA 25
| | | | | | | | | |
Oy 25 ttctgagaa 33

RESULT 32
ID PCT-US96-00845-11 STANDARD; DNA; UNC; 32 BP.
AC xxxxxx
DE Sequence 11, Application PC/TUS9600845
CC Sequence 11, Application PC/TUS9600845
CC GENERAL INFORMATION:
CC APPLICANT: Fox Chase, Cancer Center
CC APPLICANT: Golemis, Erica A.
CC APPLICANT: Khazak, Vladimir
CC APPLICANT: Estojak, Joanne
CC TITLE OF INVENTION: Nucleic Acid Encoding a
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dann, Dorfmann, Herrell and Skillman
CC STREET: 1601 Market Street Suite 720
CC CITY: Philadelphia
CC STATE: PA
CC COUNTRY: USA
CC ZIP: 19103-2307
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/00845
CC FILING DATE: 17-JAN-1996
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/376,157
CC FILING DATE: 20-JAN-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Reed, Janet E.
CC REGISTRATION NUMBER: 36,252
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (215) 563-4100
CC TELEFAX: (215) 563-4044
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 32 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 32 BP; 12 A; 4 C; 8 G; 8 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 18 TTCTGAGAA 26
| | | | | | | | | |
Cp 33 ttctgagaa 25
RESULT 33
ID US-08-376-157B-11 STANDARD; DNA; UNC; 32 BP.
AC xxxxxx
DE Sequence 11, Application US/08376157B
CC Sequence 11, Application US/08376157B
CC Patent No. 570717
CC GENERAL INFORMATION:
CC APPLICANT: Fox Chase, Cancer Center
CC APPLICANT: Golemis, Erica A.
CC APPLICANT: Khazak, Vladimir
CC APPLICANT: Estojak, Joanne
CC TITLE OF INVENTION: Nucleic Acid Encoding a
CC TITLE OF INVENTION: Stress-Responsive Subunit of Human RNA Polymerase II

Db 9 TTCTCAGAA 17
|||||
Cc 33 ttctcagaa 25

RESULT 36
ID US-07-938-084-3 STANDARD; DNA; UNC; 40 BP.
AC xxxxxx

DE Sequence 3, Application US/07938084
CC Sequence 3, Application US/07938084
CC Patent No. 5464945

CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsn, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828

CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 40 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 40 BP; 9 A; 10 C; 7 G; 14 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 TTCTCAGAA 17
|||||
Cc 33 ttctcagaa 25

RESULT 37
ID US-08-375-241-30 STANDARD; DNA; UNC; 41 BP.
AC xxxxxx

DE Sequence 30, Application US/08375241
CC Sequence 30, Application US/08375241
CC Patent No. 5648461

CC GENERAL INFORMATION:
CC APPLICANT: Parodos, Kyriaki
CC APPLICANT: McCarly, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC TITLE OF INVENTION: Shigella
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/375,241
CC FILING DATE:

CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: GTR90-04
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-9540
CC TELEFAX: 617-861-6240
CC INFORMATION FOR SEQ ID NO: 30:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 41 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 41 BP; 11 A; 11 C; 8 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 20 TTCTGAGAA 28
|||||
Cc 25 ttctgagaa 33

RESULT 38
ID PCT-US92-06617A-30 STANDARD; DNA; UNC; 41 BP.
AC xxxxxx

DE Sequence 30, Application PCT/US9206617A
CC Sequence 30, Application PCT/US9206617A
CC GENERAL INFORMATION:

CC APPLICANT: Parodos, Kyriaki
CC APPLICANT: McCarly, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC TITLE OF INVENTION: Shigella
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Amoco Corporation
CC STREET: 200 East Randolph Drive, P.O. Box 87703
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: U.S.A.
CC ZIP: 60680

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/06617A
CC FILING DATE: 19920728

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Galloway, Norval B.
CC REGISTRATION NUMBER: 33,595
CC REFERENCE/DOCKET NUMBER: GTR90-04 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312-856-7180
CC TELEFAX: 312-856-4972
CC INFORMATION FOR SEQ ID NO: 30:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 41 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
SQ SEQUENCE 41 BP; 11 A; 11 C; 8 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 20 TTCTCAGAA 28
OY 25 ttctcagaa 33

RESULT 39
ID PCT-US92-06617A-18 STANDARD; DNA; UNC; 41 BP.
AC xxxxxx

DE Sequence 18, Application PC/TUS9206617A
CC Sequence 18, Application PC/TUS9206617A
CC GENERAL INFORMATION:
CC APPLICANT: Pados, Kyriaki
CC APPLICANT: McCarthy, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Amoco Corporation
CC STREET: 200 East Randolph Drive, P.O. Box 87703
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: U.S.A.
CC ZIP: 60680
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/06617A
CC FILING DATE: 1992/07/28
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Galloway, Norval B.
CC REGISTRATION NUMBER: 33,595
CC REFERENCE/DOCKET NUMBER: GTR90-04 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312-856-7180
CC TELEFAX: 312-856-4972
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 41 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
SQ SEQUENCE 41 BP; 11 A; 8 C; 11 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 14 TTCTCAGAA 22

CP 33 ttctcagaa 25

RESULT 40
ID US-08-375-241-18 STANDARD; DNA; UNC; 41 BP.
AC xxxxxx

DE Sequence 18, Application US/08375241
CC Sequence 18, Application US/08375241
CC Patent No. 5648481
CC GENERAL INFORMATION:
CC APPLICANT: Pados, Kyriaki
CC APPLICANT: McCarthy, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millita Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/375,241
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: GTR90-04
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-9540
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 41 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
SQ SEQUENCE 41 BP; 11 A; 8 C; 11 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 14 TTCTCAGAA 22
CP 33 ttctcagaa 25

RESULT 41
ID US-08-564-955-17 STANDARD; DNA; UNC; 42 BP.
AC xxxxxx

DE Sequence 17, Application US/08564955
CC Sequence 17, Application US/08564955
CC Patent No. 5811238
CC GENERAL INFORMATION:
CC APPLICANT: STEWART, WILLEM P. C.
CC APPLICANT: CRAWFORD, ANDREAS M.
CC TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES
CC TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AN

CC NUMBER OF SEQUENCES: 67
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW
CC STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: U.S.A.
CC ZIP: 94111-3834
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/564,955
CC FILING DATE: 30-NOV-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/198,431
CC FILING DATE: 17-FEB-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/537,874
CC FILING DATE: 30-OCT-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02126
CC FILING DATE: 17-FEB-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: DUNN, TRACY J.
CC REGISTRATION NUMBER: 34,587
CC REFERENCE/DOCKET NUMBER: 165283-014611US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 326-2400
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 42 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
SQ SEQUENCE 42 BP; 14 A; 9 C; 7 G; 12 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 TTCACGAA 11
Cp 33 ttctcagaa 25

RESULT 42
ID US-08-537-874-17 STANDARD; DNA; UNC; 42 BP.
AC xxxxxx
DT
DE Sequence 17, Application US/08537874
CC Sequence 17, Application US/08537874
CC Patent No. 5830721
CC GENERAL INFORMATION:
CC APPLICANT: Stemmer, Willem P.C.
CC TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation
CC NUMBER OF SEQUENCES: 62
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/537,874
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP PCT/US95/02126
CC FILING DATE: 17-FEB-1995
CC APPLICATION NUMBER: US 08/198,431
CC FILING DATE: 17-FEB-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Liebeschuetz, Joe
CC REGISTRATION NUMBER: 37,505
CC REFERENCE/DOCKET NUMBER: 018097-014610
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-576-0200
CC TELEFAX: 415-576-0300
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 42 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
SQ SEQUENCE 42 BP; 14 A; 9 C; 7 G; 12 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 TTCACGAA 11
Cp 33 ttctcagaa 25

RESULT 43
ID US-08-198-431-17 STANDARD; DNA; UNC; 42 BP.
AC xxxxxx
DT
DE Sequence 17, Application US/08198431
CC Sequence 17, Application US/08198431
CC Patent No. 5605793
CC GENERAL INFORMATION:
CC APPLICANT: Stemmer, Willem P.C.
CC TITLE OF INVENTION: Methods for In Vitro Recombination
CC NUMBER OF SEQUENCES: 49
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Burns, Doane, Swecker & Mathis
CC STREET: 699 Prince Street
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22313
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/198,431
CC FILING DATE: 17-FEB-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mool, Leslie
CC REGISTRATION NUMBER: 37,047
CC REFERENCE/DOCKET NUMBER: 000324-003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-854-7400
CC TELEFAX: 415-854-8275
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 42 base pairs
CC TYPE: nucleic acid

CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (oligonucleotide)
SQ SEQUENCE 42 BP; 14 A; 9 C; 7 G; 12 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 42;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 TTCACGAA 11
|||||
Cp 33 ttctcagaa 25

RESULT 44
ID US-08-431-527A-1 STANDARD; DNA; UNC; 54 BP.
AC xxxxxx

DE Sequence 1, Application US/08431527A
CC Sequence 1, Application US/08431527A
CC Patent No. 5843650
CC GENERAL INFORMATION:
CC APPLICANT: David Segev
CC TITLE OF INVENTION: No. 5843650-enzymatic method for detecting nucleic acid seq
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Mark M. Friedman c/o Robert Sheinbein
CC STREET: 2940 Birchtree space lane
CC CITY: Silver Spring
CC STATE: Maryland
CC COUNTRY: United States of America
CC ZIP: 20906

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
CC COMPUTER: twinhead slimnote-890TX
CC OPERATING SYSTEM: MS DOS version 6.2,
CC OPERATING SYSTEM: Windows version 3.11
CC SOFTWARE: word for windows version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/431,527A
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Friedmam, Mark M.
CC REGISTRATION NUMBER: 33,883
CC REFERENCE/DOCKET NUMBER: 128/8
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 972-3-6938541
CC TELEFAX: 972-3-6938342
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 54
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 54 BP; 14 A; 14 C; 12 G; 14 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 43 TTCACGAA 51
|||||
Cp 33 ttctcagaa 25

RESULT 45
ID US-08-758-306-558 STANDARD; DNA; UNC; 54 BP.
AC xxxxxx

DE Sequence 558, Application US/08758306
CC Sequence 558, Application US/08758306
CC Patent No. 3807743
CC GENERAL INFORMATION:
CC APPLICANT: Stinchcomb, Dan T.
CC APPLICANT: McSwiggen, James A.
CC TITLE OF INVENTION: METHOD AND REAGENT FOR THE
CC TITLE OF INVENTION: TREATMENT OF DISEASES
CC TITLE OF INVENTION: ASSOCIATED WITH
CC TITLE OF INVENTION: INTERLEUKIN-2 RECEPTOR
CC NUMBER OF SEQUENCES: 1379
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lyon & Lyon
CC STREET: 633 West Fifth Street
CC STREET: Suite 4700
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: U.S.A.
CC ZIP: 90071-2066

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC MEDIUM TYPE: storage
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: IBM P.C. DOS 5.0
CC SOFTWARE: FastSeq version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/758,306
CC FILING DATE: December 3, 1996
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Warburg, Richard J.
CC REGISTRATION NUMBER: 32,327
CC REFERENCE/DOCKET NUMBER: 212/132
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 489-1600
CC TELEFAX: (213) 955-0440
CC TELEX: 67-3510
CC INFORMATION FOR SEQ ID NO: 558:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 54 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 54 BP; 17 A; 11 C; 11 G; 0 T; 15 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 54;
Best Local Similarity 66.7%; Pred. No. 3,96e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 4 UUCUCGAA 12
:::|||||
Cp 33 ttctcagaa 25

RESULT 46
ID US-08-390-850-1036 STANDARD; DNA; UNC; 54 BP.
AC xxxxxx

DE Sequence 1036, Application US/08390850
CC Sequence 1036, Application US/08390850
CC Patent No. 5612215
CC GENERAL INFORMATION:
CC APPLICANT: Draper, Kenneth G.
CC APPLICANT: Pavco, Pamela
CC APPLICANT: McSwiggen, James
CC APPLICANT: Gustofson, John
CC APPLICANT: Stinchcomb, Dan T.
CC TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT

CC TITLE OF INVENTION: OF ARTHRITIC CONDITIONS
CC NUMBER OF SEQUENCES: 1151
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lyon & Lyon
CC STREET: 633 West Fifth Street
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: U.S.A.
CC ZIP: 90071
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC MEDIUM TYPE: storage
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: IBM P.C. DOS 5.0
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/390,850
CC FILING DATE: February 17, 1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/354,920
CC FILING DATE: December 13, 1994
CC APPLICATION NUMBER: 08/152,487
CC FILING DATE: No. 5612215ember 12, 1993
CC APPLICATION NUMBER: 07/989,848
CC FILING DATE: December 7, 1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Warburg, Richard
CC REGISTRATION NUMBER: 32,327
CC REFERENCE/DOCKET NUMBER: 211/084
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 489-1600
CC TELEFAX: (213) 955-0440
CC TEXT: 67-3510
CC INFORMATION FOR SEQ ID NO: 1036:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 54 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 54 BP; 18 A; 12 C; 12 G; 0 T; 12 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 54;
Best Local Similarity 66.7%; Pred. No. 3.96e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 4 UUCUCAGAA 12
Cp 33 ttctcagaa 25
RESULT 47
ID US-08-435-634-1036 STANDARD; DNA; UNC; 54 BP.
AC xxxxxx
DE Sequence 1036, Application US/08435634
CC Sequence 1036, Application US/08435634
CC Patent No. 5731295
CC GENERAL INFORMATION:
CC APPLICANT: Draper, Kenneth G.
CC APPLICANT: Pavco, Pamela
CC APPLICANT: McSwigen, James
CC APPLICANT: Gustofson, John
CC APPLICANT: Stinchcomb, Dan T.
CC TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT
CC TITLE OF INVENTION: OF ARTHRITIC CONDITIONS
CC NUMBER OF SEQUENCES: 1151
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lyon & Lyon
CC STREET: 633 West Fifth Street
CC CITY: Los Angeles
CC STATE: California

CC COUNTRY: U.S.A.
CC ZIP: 90071
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC MEDIUM TYPE: storage
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: IBM P.C. DOS 5.0
CC SOFTWARE: FastSeq Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/435,634
CC FILING DATE: 05-MAY-1995
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/390,850
CC FILING DATE: February 17, 1995
CC APPLICATION NUMBER: 08/354,920
CC FILING DATE: December 13, 1994
CC APPLICATION NUMBER: 08/152,487
CC FILING DATE: No. 5731295ember 12, 1993
CC APPLICATION NUMBER: 07/989,848
CC FILING DATE: December 7, 1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Warburg, Richard
CC REGISTRATION NUMBER: 32,327
CC REFERENCE/DOCKET NUMBER: 211/084
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 489-1600
CC TELEFAX: (213) 955-0440
CC TEXT: 67-3510
CC INFORMATION FOR SEQ ID NO: 1036:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 54 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 54 BP; 18 A; 12 C; 12 G; 0 T; 12 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 54;
Best Local Similarity 66.7%; Pred. No. 3.96e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 4 UUCUCAGAA 12
Cp 33 ttctcagaa 25
RESULT 48
ID US-08-592-029-6 STANDARD; DNA; UNC; 54 BP.
AC xxxxxx
DE Sequence 6, Application US/08592029
CC Sequence 6, Application US/08592029
CC Patent No. 5763196
CC GENERAL INFORMATION:
CC APPLICANT: POWELL, MICHAEL J.
CC APPLICANT: KHANNA, PYARE
CC APPLICANT: KINENFELTER, DAVID
CC APPLICANT: EISENBEIS, SCOTT J.
CC TITLE OF INVENTION: ASSAYS USING CROSS-LINKED POLYPEPTIDE
CC TITLE OF INVENTION: FRAGMENTS OF beta-GALACTOSIDASE
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MORRISON & FOERSTER
CC STREET: 755 PAGE MILL ROAD
CC CITY: PALO ALTO
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304-1018
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/592,029
CC FILING DATE: 26-JAN-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: MONROY, GLADYS H.
CC REGISTRATION NUMBER: 32,430
CC REFERENCE/DOCKET NUMBER: 33746-20004.00
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 813-5600
CC TELEFAX: (415) 494-0792
CC TELEX: 706141 MRSNFOERS SFO
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 54 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SQ SEQUENCE 54 BP; 14 A; 14 C; 10 G; 16 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 9 TTCTCAGAA 17
CP 33 ttctcagaa 25

RESULT 49
ID US-07-938-084-9 STANDARD; DNA; UNC; 60 BP.
AC xxxxxx

DE Sequence 9, Application US/07938084
CC Sequence 9, Application US/07938084
CC Patent No. 5464945
CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsh, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 60 BP; 20 A; 10 C; 11 G; 19 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 9 TTCTCAGAA 17
CP 33 ttctcagaa 25

RESULT 50
ID PCT-US96-03940-15 STANDARD; DNA; UNC; 63 BP.
AC xxxxxx

DE Sequence 15, Application PC/TUS9603940
CC Sequence 15, Application PC/TUS9603940
CC GENERAL INFORMATION:
CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC TITLE OF INVENTION: REST Protein and DNA
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechert Price & Rhoads
CC STREET: P.O. Box 5218
CC CITY: Princeton
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 08543-5218
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
CC COMPUTER: IBM-compatible
CC OPERATING SYSTEM: DOS 5.0
CC SOFTWARE: WordPerfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03940
CC FILING DATE: March 23, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Allen Bloom
CC REGISTRATION NUMBER: 29,135
CC REFERENCE/DOCKET NUMBER: 317743-101
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214
CC TELEFAX: (609) 520-3259
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 63 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC MOLECULE TYPE: linear
CC TOPOLOGY: linear
CC MEDIUM TYPE: CDNA to mRNA
CC HYPOTHETICAL: no
CC ANTI-SENSE: no
CC ORIGINAL SOURCE:
CC ORGANISM: Human
CC CELL LINE: HELA
CC IMMEDIATE SOURCE:
CC LIBRARY: CDNA
CC PUBLICATION INFORMATION:
CC AUTHORS: Chong, Jayhong A., Tapia-Ramirez Jos, Toledo-
CC AUTHORS: Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
CC AUTHORS: Yelena M., Frohman, Michael A., Krane, Susan D., Mandel, Gail
CC TITLE: REST: A Mammalian Silencer Protein that Restricts
CC TITLE: Sodium Channel Gene Expression to Neurons
CC JOURNAL: Cell
CC VOLUME: 80
CC ISSUE:
CC PAGES:
CC DATE: March 24, 1995
CC RELEVANT RESIDUES IN SEQ ID NO: 15: FROM 1 TO 63
SQ SEQUENCE 63 BP; 18 A; 15 C; 9 G; 21 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 63;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 TTCTCAGAA 32

Cp 33 ttctcagaa 25

RESULT 51

ID US-07-982-744B-2 STANDARD: DNA: UNC: 63 BP.

AC xxxxxx

DE Sequence 2, Application US/07982744B

CC Sequence 2, Application US/07982744B

CC Patent No. 5591604

CC GENERAL INFORMATION:

CC APPLICANT: Fuchs, Patrick

CC APPLICANT: Little, Melvyn

CC APPLICANT: Breiling, Frank

CC APPLICANT: D bel, Stefan

CC TITLE OF INVENTION: Recombinant Antibodies at the surface of

CC NUMBER OF SEQUENCES: 6

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH

CC STREET: 8110 Gatehouse Road, Suite 500 East

CC CITY: Falls Church

CC STATE: Virginia

CC COUNTRY: U.S.A.

CC ZIP: 22042

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette

CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: word perfect, converted into ASCII-codes

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/982.744B

CC FILING DATE: 10-MAY-1993

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: DE P4122598.8

CC FILING DATE: 08-JUL-1991

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: PCT/EP92/01523

CC FILING DATE: 06-JUL-1992

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Svensson, Leonard R.

CC REGISTRATION NUMBER: 30.330

CC REFERENCE/DOCKET NUMBER: 1879-102PCT

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (703) 205-8000

CC TELEFAX: (703) 205-8050

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 63 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double stranded

CC TOPOLOGY: linear

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 1..63

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 2..63

CC OTHER INFORMATION: /product = "protein containing the epitope

CC OTHER INFORMATION: for the tubulin monoclonal antibody 10L1/34

CC SEQUENCE 63 BP: 19 A; 15 C; 18 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 63;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 TTCTCAGAA 48

Cp 33 ttctcagaa 25

RESULT 52

ID US-07-982-743-16 STANDARD: DNA: UNC: 63 BP.

AC xxxxxx

DE Sequence 16, Application US/07982743

CC Sequence 16, Application US/07982743

CC Patent No. 5849500

CC GENERAL INFORMATION:

CC APPLICANT: Breiling, Frank

CC APPLICANT: Little, Melvyn

CC APPLICANT: Dubel, Stefan

CC APPLICANT: Braunagel, Michael

CC APPLICANT: Klewinghaus, Iris

CC TITLE OF INVENTION: Phagemid for Antibody Screening

CC NUMBER OF SEQUENCES: 17

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Birch, Stewart, Kolasch & Birch

CC STREET: P.O. Box 747

CC CITY: Falls Church

CC STATE: Virginia

CC COUNTRY: USA

CC ZIP: 22040-0747

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/982.743

CC FILING DATE: 10-MAY-1993

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Svensson, Leonard R.

CC REGISTRATION NUMBER: 30330

CC REFERENCE/DOCKET NUMBER: 1879-101PCT

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 703-205-8000

CC TELEFAX: 703-205-8050

CC INFORMATION FOR SEQ ID NO: 16:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 63 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: both

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA

CC FEATURE:

CC NAME/KEY: -

CC LOCATION: 1..63

CC OTHER INFORMATION: /label= linker

CC OTHER INFORMATION: /note= "alternative tag-linker sequence, see Fig.

CC OTHER INFORMATION: 1(c)"

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 1..63

CC SEQUENCE 63 BP: 19 A; 15 C; 18 G; 11 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 63;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 TTCTCAGAA 48

Cp 33 ttctcagaa 25

RESULT 53

ID US-07-982-743-9 STANDARD: DNA: UNC: 66 BP.

AC xxxxxx

DT

DE Sequence 9, Application US/07982743
CC Sequence 9, Application US/07982743
CC Patent No. 5849500
CC GENERAL INFORMATION:
CC APPLICANT: Bretling, Frank
CC APPLICANT: Little, Melynn
CC APPLICANT: Dubel, Stefan
CC APPLICANT: Braunagel, Michael
CC APPLICANT: Kiewinghaus, Iris
CC TITLE OF INVENTION: Phagemid for Antibody screening
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Birch, Stewart, Kolasch & Birch
CC STREET: P.O. Box 747
CC CITY: Falls Church
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22040-0747
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/982,743
CC FILING DATE: 10-MAY-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Svensson, Leonard R.
CC REGISTRATION NUMBER: 30330
CC REFERENCE/DOCKET NUMBER: 1879-101PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-205-8000
CC TELEFAX: 703-205-8050
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 66 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: both
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA
CC FEATURE:
CC NAME/KEY: 1..66
CC LOCATION: 1..66
CC OTHER INFORMATION: /label= fragment
CC OTHER INFORMATION: /note= "tag-linker/scab boundary in plasmid pSEX
CC OTHER INFORMATION: (nucleotides 523 - 589), see Fig. 1"
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..66
CC SEQUENCE 66 BP; 21 A; 11 C; 18 G; 16 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 19 TTCTGAGAA 27
1111111111
Cc 33 ttctcgagaa 25

RESULT 54
ID 5200336-2 STANDARD; DNA; UNC; 80 BP.
AC xxxxxx
DT 01-JAN-1900
DE Patent No. 5200336.
CC Patent No. 5200336
CC APPLICANT: KONG, HUIMIN;SCHILDKRAUT, IRA
CC TITLE OF INVENTION: RESTRICTION ENDONUCLEASE OBTAINABLE FROM
CC BACILLUS COAGULANS AND A PROCESS FOR PRODUCING THE SAME
CC NUMBER OF SEQUENCES: 3
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/547,787

CC FILING DATE: 02-JUL-1990
CC SEQ ID NO:2:
CC LENGTH: 74
CC Sequence 80 BP; 18 A; 18 C; 20 G; 18 T; 6 other;
S0 Query Match 100.0%; Score 9; DB 5; Length 74;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 18 TTCTGAGAA 26
1111111111
Cc 25 ttctcgagaa 33

RESULT 55
ID US-08-477-830-5 STANDARD; DNA; UNC; 77 BP.
AC xxxxxx
DT Sequence 5, Application US/08477830
DE Sequence 5, Application US/08477830
CC Patent No. 5654151
CC GENERAL INFORMATION:
CC APPLICANT: PATRICK ALLEN
CC APPLICANT: LARRY GOLD
CC TITLE OF INVENTION: HIGH AFFINITY HIV
CC TITLE OF INVENTION: NUCLEOCAPSID NUCLEIC ACID LIGANDS
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Swanson and Bratschun, L.L.C.
CC STREET: 8400 East Prentice Avenue, Suite #200
CC CITY: Denver
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: Wordperfect 6.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/477,830
CC FILING DATE: 7-JUNE-1995
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/714,131
CC FILING DATE: 10-JUNE-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/536,428
CC FILING DATE: 11-JUNE-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/964,624
CC FILING DATE: 21-OCTOBER-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/117,991
CC FILING DATE: 8-SEPTEMBER-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/931,474
CC FILING DATE: 17-AUGUST-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/361,795
CC FILING DATE: 21-DECEMBER-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/477,172
CC FILING DATE: 19-MAY-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Diane H. McClearn
CC REGISTRATION NUMBER: 33,960
CC REFERENCE/DOCKET NUMBER: NEX 44-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (303) 793-3333
CC TELEFAX: (303) 793-3433
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 77 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: RNA
 SO SEQUENCE 77 BP; 18 A; 22 C; 20 G; 0 T; 17 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 77;
 Best Local Similarity 66.7%; Pred. No. 3,96e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 25 UUCUGAGAA 33
 :|:|||||
 QY 25 ttctgagaa 33

RESULT 56
 ID US-08-477-530-5 STANDARD; DNA; UNC; 77 BP.
 AC xxxxxx
 DT Sequence 5, Application US/08477530
 DE Sequence 5, Application US/08477530
 CC Patent No. 5635400
 CC GENERAL INFORMATION:
 CC APPLICANT: PATRICK ALLEN
 CC APPLICANT: LARRY GOLD
 CC TITLE OF INVENTION: HIGH AFFINITY HIV
 CC NUMBER OF SEQUENCES: 29
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Swanson and Bratschun, L.L.C.
 CC STREET: 8400 East Prentice Avenue, Suite #200
 CC CITY: Denver
 CC STATE: Colorado
 CC COUNTRY: USA
 CC ZIP: 80111
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
 CC OPERATING SYSTEM: MS-DOS
 CC SOFTWARE: WordPerfect 6.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/477,530
 CC FILING DATE: 7-JUNE-1995
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/714,131
 CC FILING DATE: 10-JUNE-1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/536,428
 CC FILING DATE: 11-JUNE-1990
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/964,624
 CC FILING DATE: 21-OCTOBER-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/117,991
 CC FILING DATE: 8-SEPTEMBER-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/931,474
 CC FILING DATE: 17-AUGUST-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/361,795
 CC FILING DATE: 21-DECEMBER-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/477,172
 CC FILING DATE: 19-MAY-1995
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Diane H. McLearn
 CC REGISTRATION NUMBER: 33,960
 CC REFERENCE/DOCKET NUMBER: NEX 44-2
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (303) 793-3333
 CC TELEFAX: (303) 793-3433

CC INFORMATION FOR SEQ ID NO: 5:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 77 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: RNA
 SO SEQUENCE 77 BP; 18 A; 22 C; 20 G; 0 T; 17 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 77;
 Best Local Similarity 66.7%; Pred. No. 3,96e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 25 UUCUGAGAA 33
 :|:|||||
 QY 25 ttctgagaa 33

RESULT 57
 ID US-08-477-530-5 STANDARD; DNA; UNC; 77 BP.
 AC xxxxxx
 DT Sequence 5, Application US/08477530
 DE Sequence 5, Application US/08477530
 CC Patent No. 5635615
 CC GENERAL INFORMATION:
 CC APPLICANT: PATRICK ALLEN
 CC APPLICANT: LARRY GOLD
 CC TITLE OF INVENTION: HIGH AFFINITY HIV
 CC NUMBER OF SEQUENCES: 29
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Swanson and Bratschun, L.L.C.
 CC STREET: 8400 East Prentice Avenue, Suite #200
 CC CITY: Denver
 CC STATE: Colorado
 CC COUNTRY: USA
 CC ZIP: 80111
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
 CC OPERATING SYSTEM: MS-DOS
 CC SOFTWARE: WordPerfect 6.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/477,530
 CC FILING DATE: 7-JUNE-1995
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/714,131
 CC FILING DATE: 10-JUNE-1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/536,428
 CC FILING DATE: 11-JUNE-1990
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/964,624
 CC FILING DATE: 21-OCTOBER-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/117,991
 CC FILING DATE: 8-SEPTEMBER-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/931,474
 CC FILING DATE: 17-AUGUST-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/361,795
 CC FILING DATE: 21-DECEMBER-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/477,172
 CC FILING DATE: 19-MAY-1995
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Diane H. McLearn
 CC REGISTRATION NUMBER: 33,960
 CC REFERENCE/DOCKET NUMBER: NEX 44-2
 CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (303) 793-3333
CC TELEFAX: (303) 793-3433
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 77 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: RNA
CC SEQUENCE 77 BP; 18 A; 22 C; 20 G; 0 T; 17 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 77;
Best Local Similarity 66.7%; Pred. No. 3,966+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 25 UUCUGAGAA 33
Oy 25 ttctgagaa 33

RESULT 58
ID US-08-441-430-18 STANDARD; DNA; UNC; 80 BP.
AC xxxxxx
DT Sequence 18, Application US/08441430
DE Sequence 18, Application US/08441430
CC Patent No. 5681942
CC GENERAL INFORMATION:
CC APPLICANT: Buchwald, Manuel
CC APPLICANT: Strathdee, Craig A.
CC APPLICANT: Wevrick, Rachel
CC APPLICANT: Mathew, Christopher George Porter
CC TITLE OF INVENTION: Fanconi Anemia Type C Gene
CC NUMBER OF SEQUENCES: 73
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Richard J. Polley, Esq.
CC ADDRESSEE: Klarquist, Sparkman, Campbell, Leigh &
CC ADDRESSEE: Whinston, LLP
CC STREET: 121 S.W. Salmon, Suite 1600
CC CITY: Portland
CC STATE: Oregon
CC COUNTRY: U.S.A.
CC ZIP: 97204
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Disk, 3+1/2-inch
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: MS DOS
CC SOFTWARE: WordPerfect 5.1/ASCII Text File
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/441,430
CC FILING DATE: May 15, 1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: U.S. 07/876,285
CC FILING DATE: April 29, 1992
CC APPLICATION NUMBER: U.S. 07/918,313
CC FILING DATE: July 21, 1992
CC APPLICATION NUMBER: U.S. 08/003,963
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Richard J. Polley, Esq.
CC REGISTRATION NUMBER: 28,107
CC REFERENCE/DOCKET NUMBER: 3812-42824
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (503) 226-7391
CC TELEFAX: (503) 228-9446
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 80 base pairs
CC TYPE: Nucleic Acid
CC STRANDEDNESS: Double stranded
CC TOPOLOGY: Linear
CC MOLECULE TYPE: Genomic DNA

CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC SEQUENCE 80 BP; 19 A; 15 C; 14 G; 31 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 80;
Best Local Similarity 88.9%; Pred. No. 3,966+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 27 TTYTGAGAA 35
Oy 25 ttctgagaa 33

RESULT 59
ID US-07-938-084-10 STANDARD; DNA; UNC; 80 BP.
AC xxxxxx
DT Sequence 10, Application US/07938084
DE Sequence 10, Application US/07938084
CC Patent No. 5464945
CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsh, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 80 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 80 BP; 24 A; 15 C; 16 G; 25 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 3,966+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 TTCGAGAA 17
Oy 33 ttctcagaa 25

RESULT 60
ID US-07-938-084-11 STANDARD; DNA; UNC; 100 BP.
AC xxxxxx
DT

DE Sequence 11, Application US/07938084
CC Sequence 11, Application US/07938084
CC Patent No. 5464945
CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsh, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 100 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 100 BP: 30 A; 18 C; 20 G; 32 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 TTCTCAGAA 17
| | | | | | | | | |
Cp 33 ttctcagaa 25

RESULT 61
ID US-07-938-084-4 STANDARD; DNA; UNC; 130 BP.
AC xxxxxx
DE Sequence 4, Application US/07938084
CC Sequence 4, Application US/07938084
CC Patent No. 5464945
CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsh, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 130 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 130 BP: 37 A; 23 C; 30 G; 40 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 TTCTCAGAA 17
| | | | | | | | | |
Cp 33 ttctcagaa 25

RESULT 62
ID US-07-938-084-6 STANDARD; DNA; UNC; 130 BP.
AC xxxxxx
DE Sequence 6, Application US/07938084
CC Sequence 6, Application US/07938084
CC Patent No. 5464945
CC GENERAL INFORMATION:
CC APPLICANT: Reynolds, Rebecca L.
CC APPLICANT: Walsh, P. Sean
CC TITLE OF INVENTION: A Chemiluminescent Method for the
CC TITLE OF INVENTION: Quantitation of Human DNA
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: NJ
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/938,084
CC FILING DATE: 19920828
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sias Ph.D., Stacey R.
CC REGISTRATION NUMBER: 32,630
CC REFERENCE/DOCKET NUMBER: 8669
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (510) 814-2863
CC TELEFAX: (510) 522-1285
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 130 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 130 BP; 32 A; 26 C; 28 G; 44 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 9 TTCTCAGAA 17
|||||
CP 33 ttctcagaa 25

RESULT 63

ID US-08-343-943-12 STANDARD; DNA; UNC; 132 BP.
AC xxxxxx

DE Sequence 12, Application US/08343943
CC Sequence 12, Application US/08343943
CC Patent No. 5585248

GENERAL INFORMATION:

CC APPLICANT: ASHIDA, MASAKI
CC APPLICANT: KAWABATA, TOMOHIISA
CC APPLICANT: HIRAYASU, KAZUNARI
CC APPLICANT: TSUCHIYA, MASAKAKU
CC TITLE OF INVENTION: METHOD FOR ASSAYING ACTIVITY OF
CC TITLE OF INVENTION: PROPHENOXIDASE ACTIVATING ENZYME AND APPLICATION
CC TITLE OF INVENTION: THEREOF
CC NUMBER OF SEQUENCES: 13
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
CC STREET: 2100 Pennsylvania Avenue, N.W.
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA

CC ZIP: 20037-3202

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette

CC COMPUTER: IBM Compatible

CC OPERATING SYSTEM: DOS

CC SOFTWARE: FASTSEQ Version 1.5

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/343.943

CC FILING DATE: 17-NOV-1994

CC CLASSIFICATION: 435

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: 289513

CC FILING DATE: 18-NOV-1993

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Becker, Cheryl L.

CC REGISTRATION NUMBER: 35,441

CC REFERENCE/DOCKET NUMBER: 036816

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 202-293-7060

CC TELEFAX: 202-293-7860

CC TELEX:

CC INFORMATION FOR SEQ ID NO: 12:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 132 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cDNA

CC HYPOTHETICAL: NO

CC ANTI-SENSE: NO

CC FRAGMENT TYPE:

CC ORIGINAL SOURCE:

CC NAME/KEY: Coding Sequence

CC LOCATION: 1..132

CC SEQUENCE 132 BP; 46 A; 34 C; 21 G; 31 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 81 TTCTCAGAA 89
|||||
CY 25 ttctcagaa 33

RESULT 64

ID US-08-700-575-28 STANDARD; DNA; UNC; 156 BP.
AC xxxxxx

DE Sequence 28, Application US/08700575
CC Sequence 28, Application US/08700575
CC Patent No. 5817479

GENERAL INFORMATION:

CC APPLICANT: Au-Yang, Janice
CC APPLICANT: Bandman, Olga
CC APPLICANT: Hawkins, Phillip R.
CC APPLICANT: Wilde, Craig G.
CC TITLE OF INVENTION: NOVEL HUMAN KINASE HOMOLOGS
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CA
CC COUNTRY: USA

CC ZIP: 94304

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/700.575

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: BILLINGS, LUCY J

CC REGISTRATION NUMBER: 36749

CC REFERENCE/DOCKET NUMBER: SP-100 US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 415-845-4166

CC TELEFAX: 415-845-4166

CC INFORMATION FOR SEQ ID NO: 28:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 156 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: cDNA

CC IMMEDIATE SOURCE:

CC LIBRARY: Bone Marrow

CC CLONE: 132750

CC SEQUENCE 156 BP; 40 A; 29 C; 33 G; 54 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 147 TTCTGAGAA 155
|||||
CY 25 ttctgagaa 33

RESULT 65

ID US-07-858-124-16 STANDARD; DNA; UNC; 161 BP.
AC xxxxxx

DE Sequence 16, Application US/07858124
CC Sequence 16, Application US/07858124
CC Patent No. 5427932

GENERAL INFORMATION:

CC APPLICANT: Welter, Heinz-Ulrich G
CC APPLICANT: Gray, Joe W
CC TITLE OF INVENTION: Repeat Sequence Chromosome Specific
CC TITLE OF INVENTION: Nucleic Acid Probes
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Leona L. Lauder
CC STREET: Steuart Street Tower, 18th Fl., One Market
CC STREET: Plaza
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/858,124
CC FILING DATE: 19920326
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/883,441
CC FILING DATE: 09-APR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lauder, Leona L
CC REGISTRATION NUMBER: 30863
CC REFERENCE/DOCKET NUMBER: 89-263-2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-777-9257
CC TELEFAX: 415-543-4219
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 161 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: GENOMIC DNA --human alpha satellite
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC PUBLICATION INFORMATION:
CC AUTHORS: Willard and Wray
CC TITLE: "Hierarchal order in chromosome-specific
CC TITLE: human alpha satellite DNA"
CC JOURNAL: Trends in Genet.
CC VOLUME: 3
CC ISSUE:
CC PAGES: 192-198
CC DATE: 1987
CC SEQUENCE 161 BP: 47 A; 22 C; 28 G; 47 T; 17 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 3,966+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 140 TTCGAGAA 148
CP 33 ttctcagaa 25
RESULT 66
ID US-08-398-627-43 STANDARD; DNA; UNC; 171 BP.
AC xxxxxx
DE Sequence 43, Application US/08398627
CC Sequence 43, Application US/08398627
CC Patent No. 5608149
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Stark, David M.

CC TITLE OF INVENTION: Enhanced Starch Biosynthesis
CC NUMBER OF SEQUENCES: 51
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Grace L. Bonner, Monsanto Co. B4F
CC STREET: 700 Chesterfield Parkway No. 5608149th
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/398,627
CC FILING DATE: 03-MAR-1995
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/090,523
CC FILING DATE: 12-JUL-1993
CC APPLICATION NUMBER: US 07/709663
CC FILING DATE: 07-JUN-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/539763
CC FILING DATE: 18-JUN-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bonner, Grace L.
CC REGISTRATION NUMBER: 32,963
CC REFERENCE/DOCKET NUMBER: 38-21(10559)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314) 537-7286
CC TELEFAX: (314) 537-6047
CC INFORMATION FOR SEQ ID NO: 43:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 171 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 171 BP: 48 A; 32 C; 41 G; 50 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 3,966+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2 TTCGAGAA 10
CY 25 ttctcagaa 33
RESULT 67
ID US-08-090-523-43 STANDARD; DNA; UNC; 171 BP.
AC xxxxxx
DE Sequence 43, Application US/08090523
CC Sequence 43, Application US/08090523
CC Patent No. 5498830
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Stark, David M.
CC TITLE OF INVENTION: Enhanced Starch Biosynthesis
CC NUMBER OF SEQUENCES: 51
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Grace L. Bonner, Monsanto Co. B4F
CC STREET: 700 Chesterfield Parkway No. 5498830th
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC *CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/090.523
CC FILING DATE: 19930712
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/709663
CC FILING DATE: 07-JUN-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/539763
CC FILING DATE: 18-JUN-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bonner, Grace L
CC REGISTRATION NUMBER: 32,963
CC REFERENCE/DOCKET NUMBER: 38-21(10559)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314) 537-7286
CC TELEFAX: (314) 537-6047
CC INFORMATION FOR SEQ ID NO: 43:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 171 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC SEQUENCE 171 BP: 48 A; 32 C; 41 G; 50 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2 TTCGAGAA 10
|||||
QY 25 ttctgagaa 33
RESULT 68
ID US-07-858-124-20 STANDARD; DNA: UNC; 171 BP.
AC xxxxxx
DE Sequence 20, Application US/07858124
CC Sequence 20, Application US/07858124
CC Patent No. 5427932
CC GENERAL INFORMATION:
CC APPLICANT: Welter, Heinz-Ulrich G
CC APPLICANT: Gray, Joe W
CC TITLE OF INVENTION: Repeat Sequence Chromosome Specific
CC TITLE OF INVENTION: Nucleic Acid Probes
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Leona L. Lauder
CC STREET: Stewart Street Tower, 18th Fl., One Market
CC STREET: Plaza
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/858.124
CC FILING DATE: 19920326
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/683.441
CC FILING DATE: 09-APR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lauder, Leona L
CC

CC REGISTRATION NUMBER: 30863
CC REFERENCE/DOCKET NUMBER: 89-263-2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-777-9257
CC TELEFAX: 415-543-4219
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 171 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: GENOMIC DNA -- chromosome 12-specific
CC MOLECULE TYPE: aliphoid monomer sequence
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC PUBLICATION INFORMATION:
CC AUTHORS: Baldini et al.
CC TITLE: "High-resolution Alu banding of human
CC TITLE: chromosomes with biotinylated Alu-PCR products:
CC TITLE: use for in situ hybridization mapping" (Abstract)
CC JOURNAL: Am. J. Hum. Genet
CC VOLUME: 46 (Suppl.)
CC PAGES: A87
CC DATE: 1990
CC SEQUENCE 171 BP: 54 A; 29 C; 37 G; 51 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 141 TTCGAGAA 149
|||||
Cp 33 ttctcagaa 25
RESULT 69
ID US-07-858-124-22 STANDARD; DNA: UNC; 171 BP.
AC xxxxxx
DE Sequence 22, Application US/07858124
CC Sequence 22, Application US/07858124
CC Patent No. 5427932
CC GENERAL INFORMATION:
CC APPLICANT: Welter, Heinz-Ulrich G
CC APPLICANT: Gray, Joe W
CC TITLE OF INVENTION: Repeat Sequence Chromosome Specific
CC TITLE OF INVENTION: Nucleic Acid Probes
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Leona L. Lauder
CC STREET: Stewart Street Tower, 18th Fl., One Market
CC STREET: Plaza
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/858.124
CC FILING DATE: 19920326
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/683.441
CC FILING DATE: 09-APR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lauder, Leona L
CC REGISTRATION NUMBER: 30863
CC REFERENCE/DOCKET NUMBER: 89-263-2
CC TELECOMMUNICATION INFORMATION:
CC

CC TELEPHONE: 415-777-9257
 CC TELEFAX: 415-543-4219
 CC INFORMATION FOR SEQ ID NO: 22:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 171 base pairs
 CC TYPE: NUCLEIC ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: GENOMIC DNA -- chromosome 12-specific
 CC MOLECULE TYPE: alphaoid monomer sequence
 CC HYPOTHETICAL: NO
 CC ANTI-SENSE: NO
 CC PUBLICATION INFORMATION:
 CC AUTHORS: Baldwin et al.
 CC TITLE: "High-resolution Alu banding of human
 CC TITLE: chromosomes with biotinylated Alu-PCR products:
 CC JOURNAL: Am. J. Hum. Genet
 CC VOLUME: 46 (Suppl.)
 CC PAGES: A87
 CC DATE: 1990
 CC
 CC SEQUENCE 171 BP: 57 A; 34 C; 30 G; 50 T; 0 OTHER.
 CC
 CC Query Match 100.0%; Score 9; DB 1; Length 171;
 CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 141 TTCTCAGAA 149
 CP 33 ttctcagaa 25
 AC xxxxxx
 DE Sequence 21, Application US/08151477A
 DE Sequence 21, Application US/08151477A
 CC Patent No. 5830644
 CC GENERAL INFORMATION:
 CC APPLICANT: Michael D. West
 CC APPLICANT: Jerry W. Shay
 CC APPLICANT: Woodring E. Wright
 CC APPLICANT: Elizabeth Blackburn
 CC APPLICANT: Nam Woo Kim
 CC APPLICANT: Calvin B. Harley
 CC APPLICANT: Scott L. Weinlich
 CC APPLICANT: Catherine Strahl
 CC APPLICANT: Michael J. McEachern
 CC APPLICANT: Homayoun Vaziri
 CC TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF
 CC TITLE OF INVENTION: CONDITIONS RELATED TO TELOMERE
 CC TITLE OF INVENTION: LENGTH AND/OR TELOMERASE ACTIVITY
 CC NUMBER OF SEQUENCES: 58
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lyon & Lyon
 CC STREET: 633 West Fifth Street
 CC STREET: Suite 4700
 CC CITY: Los Angeles
 CC STATE: California
 CC COUNTRY: U.S.A.
 CC ZIP: 90071
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 CC MEDIUM TYPE: storage
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: IBM P.C. DOS 5.0
 CC SOFTWARE: Fastseq Version 1.5
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/151,477A
 CC FILING DATE: No. 5830644ember 12, 1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/038,766

CC FILING DATE: March 24, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warburg, Richard
 CC REGISTRATION NUMBER: 32,327
 CC REFERENCE/DOCKET NUMBER: 202/189
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (213) 489-1600
 CC TELEFAX: (213) 955-0440
 CC TELEEX: 67-3510
 CC INFORMATION FOR SEQ ID NO: 21:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 192
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC
 CC SEQUENCE 192 BP: 58 A; 49 C; 33 G; 52 T; 0 OTHER.
 CC
 CC Query Match 100.0%; Score 9; DB 3; Length 192;
 CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 41 TTCTCAGAA 49
 CP 33 ttctcagaa 25
 AC xxxxxx
 DE Sequence 37, Application US/08060952C
 DE Sequence 37, Application US/08060952C
 CC Patent No. 5695932
 CC GENERAL INFORMATION:
 CC APPLICANT: Michael D. West
 CC APPLICANT: Jerry W. Shay
 CC APPLICANT: Woodring E. Wright
 CC APPLICANT: Elizabeth Blackburn
 CC TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF CONDITIONS
 CC TITLE OF INVENTION: RELATED TO TELOMERE LENGTH AND/OR
 CC TITLE OF INVENTION: TELOMERASE ACTIVITY
 CC NUMBER OF SEQUENCES: 57
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lyon & Lyon
 CC STREET: 633 West Fifth Street
 CC STREET: Suite 4700
 CC CITY: Los Angeles
 CC STATE: California
 CC COUNTRY: U.S.A.
 CC ZIP: 90071-2066
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 CC MEDIUM TYPE: storage
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: IBM P.C. DOS 5.0
 CC SOFTWARE: Word Perfect 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/060,952C
 CC FILING DATE: May 13, 1993
 CC CLASSIFICATION: 514
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/882,438
 CC FILING DATE: May 13, 1992
 CC APPLICATION NUMBER: 08/038,766
 CC FILING DATE: March 24, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warburg, Richard J.
 CC REGISTRATION NUMBER: 32,327
 CC REFERENCE/DOCKET NUMBER: 202/045
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (213) 489-1600
 CC TELEFAX: (213) 955-0440
 CC TELEEX: 67-3510

CC INFORMATION FOR SEQ ID NO: 37:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 192
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SQ SEQUENCE 192 BP; 58 A; 49 C; 33 G; 52 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 41 TTCTCAGAA 49
CP 33 ttctcagaa 25

RESULT 72
ID US-08-153-051B-21 STANDARD; DNA; UNC; 192 BP.
AC xxxxxx

DE Sequence 21, Application US/08153051B
CC Sequence 21, Application US/08153051B
CC Patent No. 5645986
CC GENERAL INFORMATION:
CC APPLICANT: Michael D. West
CC APPLICANT: Jerry W. Shay
CC APPLICANT: Woodring E. Wright
CC APPLICANT: Elizabeth Blackburn
CC APPLICANT: Nam Woo Kim
CC APPLICANT: Calvin B. Harley
CC APPLICANT: Scott L. Weinrich
CC APPLICANT: Catherine Strahl
CC APPLICANT: Michael J. McEachern
CC APPLICANT: Homayoun Vaziri
CC TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF
CC TITLE OF INVENTION: CONDITIONS RELATED TO TELOMERE
CC TITLE OF INVENTION: LENGTH AND/OR TELOMERASE ACTIVITY
CC NUMBER OF SEQUENCES: 58
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lyon & Lyon
CC STREET: 633 West Fifth Street
CC STREET: Suite 4700
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: U.S.A.
CC ZIP: 90071
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 MB
CC MEDIUM TYPE: storage
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: IBM P.C. DOS 5.0
CC SOFTWARE: FASTSEQ Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/153,051B
CC FILING DATE: No. 5645986ember 12, 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/038,766
CC FILING DATE: March 24, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Warburg, Richard
CC REGISTRATION NUMBER: 32,327
CC REFERENCE/DOCKET NUMBER: 204/195
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 489-1600
CC TELEFAX: (213) 955-0440
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 192
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

SQ SEQUENCE 192 BP; 58 A; 49 C; 33 G; 52 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 41 TTCTCAGAA 49
CP 33 ttctcagaa 25

RESULT 73
ID US-07-858-124-17 STANDARD; DNA; UNC; 197 BP.
AC xxxxxx

DE Sequence 17, Application US/07858124
CC Sequence 17, Application US/07858124
CC Patent No. 5427932
CC GENERAL INFORMATION:
CC APPLICANT: Weter, Heinz-Ulrich G
CC APPLICANT: Gray, Joe W
CC TITLE OF INVENTION: Repeat Sequence Chromosome Specific
CC TITLE OF INVENTION: Nucleic Acid Probes
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Leona L. Lauder
CC STREET: Stewart Street Tower, 18th Fl., One Market
CC STREET: Plaza
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/858,124
CC FILING DATE: 19920326
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/683,441
CC FILING DATE: 09-APR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lauder, Leona L
CC REGISTRATION NUMBER: 30863
CC REFERENCE/DOCKET NUMBER: 89-263-2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-777-9257
CC TELEFAX: 415-543-4219
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 197 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: GENOMIC DNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
SQ SEQUENCE 197 BP; 61 A; 37 C; 43 G; 56 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 158 TTCTCAGAA 166
CP 33 ttctcagaa 25

RESULT 74
ID US-07-858-124-18 STANDARD; DNA; UNC; 197 BP.

AC xxxxxx
DE Sequence 18, Application US/07858124
CC Sequence 18, Application US/07858124
CC Patent No. 5427932
CC GENERAL INFORMATION:
CC APPLICANT: Weier, Heinz-Ulrich G
CC APPLICANT: Gray, Joe W
CC TITLE OF INVENTION: Repeat Sequence Chromosome Specific
CC TITLE OF INVENTION: Nucleic Acid Probes
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Leona L. Lauder
CC STREET: Stewart Street Tower, 18th Fl., One Market
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/858,124
CC FILING DATE: 19920326
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/683,441
CC FILING DATE: 09-APR-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lauder, Leona L
CC REGISTRATION NUMBER: 30863
CC REFERENCE/DOCKET NUMBER: 89-263-2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-777-9257
CC TELEFAX: 415-543-4219
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 197 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: GENOMIC DNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC SEQUENCE 197 BP: 60 A; 35 C; 42 G; 60 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 158 TTCACGAA 166
|||||
Cp 33 ttctcagaa 25
RESULT 75
ID PCT-US95-02576-29 STANDARD: DNA; UNC: 261 BP.
AC xxxxxx
DE Sequence 29, Application PC/TUS9502576
CC Sequence 29, Application PC/TUS9502576
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
CC NUMBER OF SEQUENCES: 65
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston

CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02576
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/205,697
CC FILING DATE: 02-Mar-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mandragoras, Amy E.
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: BWI-120CPPC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 29:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 261 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..135
CC SEQUENCE 261 BP: 97 A; 52 C; 59 G; 53 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 201 TTCACGAA 209
|||||
Cp 33 ttctcagaa 25
RESULT 76
ID US-08-928-613-3 STANDARD: DNA; UNC: 267 BP.
AC xxxxxx
DE Sequence 3, Application US/08928613
CC Sequence 3, Application US/08928613
CC Patent No. 5840562
CC GENERAL INFORMATION:
CC APPLICANT: Diep, Dinh
CC APPLICANT: Braxton, Scott M.
CC APPLICANT: Deleageane, Angelo M.
CC TITLE OF INVENTION: A NOVEL HUMAN CYSTEINE PROTEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/928,613
CC FILING DATE:
CC CLASSIFICATION: 524
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/567,506

CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Luther, Barbara J.
CC + REGISTRATION NUMBER: 33954
CC REFERENCE/DOCKET NUMBER: PF-0048 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-852-0195
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 267 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC ORIGINAL SOURCE:
CC HAPLOTYPE: U937
CC IMMEDIATE SOURCE:
CC LIBRARY: 001098
CC
SQ SEQUENCE 267 BP; 75 A; 40 C; 57 G; 95 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 96 TTCGAGAA 104
OY 25 ttctgagaa 33

RESULT 77
ID US-08-406-248-5 STANDARD; DNA; UNC; 297 BP.
AC xxxxxx
DE Sequence 5, Application US/08406248
CC Sequence 5, Application US/08406248
CC Patent No. 5736318
CC GENERAL INFORMATION:
CC APPLICANT: Munger, Karl
CC TITLE OF INVENTION: METHOD AND KIT FOR EVALUATING
CC TITLE OF INVENTION: TRANSFORMED CELLS
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ann-Louise Kerner, Ph.D., Lappin & Kusmer
CC STREET: 200 State Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/406,248
CC FILING DATE:
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: McDaniel, Patricia A.
CC REGISTRATION NUMBER: 33,194
CC REFERENCE/DOCKET NUMBER: HA2-011
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-330-1300
CC TELEFAX: 617-330-1311
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 297 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)

CC ORIGINAL SOURCE:
CC ORGANISM: Human papillomavirus
CC STRAIN: HPV-16
CC IMMEDIATE SOURCE:
CC CLONE: E7
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..297
CC SEQUENCE 297 BP; 93 A; 61 C; 67 G; 76 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 282 TTCGAGAA 290
CP 33 ttctcagaa 25

RESULT 78
ID US-08-171-385-1 STANDARD; DNA; UNC; 331 BP.
AC xxxxxx
DE Sequence 1, Application US/08171385
CC Sequence 1, Application US/08171385
CC Patent No. 5527884
CC GENERAL INFORMATION:
CC APPLICANT: Mary E. Russell
CC APPLICANT: Ulrike Utans
CC TITLE OF INVENTION: Mediators of Chronic Allograft
CC TITLE OF INVENTION: Rejection
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 MB
CC COMPUTER: IBM PS/2 Model 502 or 555X
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: Wordperfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/171,385
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fraser, Janis K.
CC REGISTRATION NUMBER: 34,819
CC REFERENCE/DOCKET NUMBER: 05433/006001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200134
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 331
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC
SQ SEQUENCE 331 BP; 79 A; 87 C; 65 G; 100 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 171 TTCGAGAA 179
IIIIIIIIII

Qy 25 ttctcgagaa 33

RESULT 79
ID US-08-470-720-1 STANDARD; DNA; UNC; 342 BP.
AC xxxxxx
DT

DE Sequence 1, Application US/08470720
CC Sequence 1, Application US/08470720
CC Patent No. 5824798
CC GENERAL INFORMATION:
CC APPLICANT: HOFVANDER, Per
CC APPLICANT: PERSON, Per T
CC APPLICANT: WIKSTROM, Olle
CC APPLICANT: TALBERG, Anne
CC TITLE OF INVENTION: GENETICALLY ENGINEERED MODIFICATION OF
CC TITLE OF INVENTION: POTATO TO FORM AMYLOPECTIN-TYPE STARCH
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Burns, Doane, Swecker & Mathis
CC STREET: George Mason Bldg., Washington & Prince Sts.
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: United States
CC ZIP: 22313-1404
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/470,720
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/070,455
CC FILING DATE: 09-JUN-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Crane-Feury, Sharon E
CC REGISTRATION NUMBER: 36,113
CC REFERENCE/DOCKET NUMBER: 003300-293
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 836-6620
CC TELEFAX: (703) 836-2021
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 342 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 217..342
CC SEQUENCE 342 BP; 85 A; 78 C; 54 G; 125 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 342;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 132 TTCTCAGAA 140
|||||
Cp 33 ttctcgagaa 25

RESULT 80
ID US-08-917-456-3 STANDARD; DNA; UNC; 363 BP.
AC xxxxxx
DT
DE Sequence 3, Application US/08917456
CC Sequence 3, Application US/08917456
CC Patent No. 5866368
CC GENERAL INFORMATION:

CC APPLICANT: Greenwood, Rebecca C.
CC APPLICANT: Gentry, Daniel R.
CC TITLE OF INVENTION: NOVEL grea
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechart Price & Rhoads
CC STREET: 4000 Bell Atlantic Tower, 1717 Arch stre
CC CITY: Philadelphia
CC STATE: PA
CC COUNTRY: US
CC ZIP: 19103
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/917,456
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dickinson, Todd O
CC REGISTRATION NUMBER: 28,354
CC REFERENCE/DOCKET NUMBER: GM10046
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-994-2252
CC TELEFAX: 215-994-2222
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 363 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC SEQUENCE 363 BP; 139 A; 57 C; 72 G; 95 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 363;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 194 TTCTGAGAA 202
|||||
Qy 25 ttctcgagaa 33

RESULT 81
ID US-08-743-200-11 STANDARD; DNA; UNC; 375 BP.
AC xxxxxx
DT
DE Sequence 11, Application US/08743200
CC Sequence 11, Application US/08743200
CC Patent No. 5861260
CC GENERAL INFORMATION:
CC APPLICANT: Doxsey, Stephen J.
CC TITLE OF INVENTION: DIAGNOSTIC METHODS FOR SCREENING
CC TITLE OF INVENTION: PATIENTS FOR SCLERODERMA
CC NUMBER OF SEQUENCES: 36
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: US
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ Version 2.0
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/743,200
CC FILING DATE: 05-NOV-1996
CC PRIOR APPLICATION DATA:
CC .APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fasse, J. Peter
CC REGISTRATION NUMBER: 32,983
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-542-5070
CC TELEFAX: 617-542-8906
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 375 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: Coding Sequence
CC LOCATION: 2...373
CC SEQUENCE 375 BP; 131 A; 83 C; 88 G; 73 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 3; Length 375;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 236 TTCGAGAA 244
|||||
Cp 33 ttctcagaa 25
RESULT 82
ID US-08-438-123-15 STANDARD; DNA; UNC; 464 BP.
AC xxxxxx
DT
DE Sequence 15, Application US/08438123
CC Sequence 15, Application US/08438123
CC Patent No. 5552293
CC GENERAL INFORMATION:
CC APPLICANT: Lindholm et al
CC TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lowe, Price, Leblanc & Becker
CC STREET: Suite 300, 99 Canal Center Plaza
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22314
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: DOS Text File
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/438,123
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/906,350
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: J.G. Mullins
CC REGISTRATION NUMBER: 33073
CC REFERENCE/DOCKET NUMBER: 149-011
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703 684 1111
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 464 base pairs
CC TYPE: Nucleic acid

CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 464 BP; 98 A; 109 C; 130 G; 127 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 464;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 34 TTCGAGAA 42
|||||
Cp 33 ttctcagaa 25
RESULT 83
ID US-08-009-261-4 STANDARD; DNA; UNC; 492 BP.
AC xxxxxx
DT
DE Sequence 4, Application US/08009261
CC Sequence 4, Application US/08009261
CC Patent No. 5350672
CC GENERAL INFORMATION:
CC APPLICANT: Oberst, Richard D.
CC APPLICANT: Gwaltney, Sharon M.
CC APPLICANT: Hays, Michael P.
CC TITLE OF INVENTION: Specific DNA Primers and Method to Use
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Tilton, Fallon, Lungmus and Chestnut
CC STREET: 100 S. Wacker Drive, Suite 960
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606-4002
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/009,261
CC FILING DATE: 19930122
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fentress, Susan B.
CC REGISTRATION NUMBER: 31,327
CC REFERENCE/DOCKET NUMBER: 92061A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312)456-8000
CC TELEFAX: (312)456-7776
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 492 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 492 BP; 128 A; 118 C; 77 G; 169 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 492;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 189 TTCGAGAA 197
|||||
Cp 25 ttctcagaa 33
RESULT 84
ID PCT-US93-10443-6 STANDARD; DNA; UNC; 495 BP.
AC xxxxxx
DT

DE Sequence 6, Application PC/TUS9310443
CC Sequence 6, Application PC/TUS9310443
CC GENERAL INFORMATION:
CC APPLICANT: David D. Moore
CC APPLICANT: Jae W. Lee
CC TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
CC TITLE OF INVENTION: INTERACTING POLYPEPTIDES AND
CC TITLE OF INVENTION: RELATED MOLECULES AND METHODS
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: WordPerfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/10443
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/969,136
CC FILING DATE: October 30, 1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Paul T. Clark
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 00786/099002
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC SEQUENCE 495 BP: 152 A; 130 C; 98 G; 114 T; 1 OTHER.
SQ
Query Match 100.0%; Score 9; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 410 TTCTCAGAA 418
CP 33 ttctcagaa 25
RESULT 85
ID US-08-222-719-6 STANDARD; DNA; UNC; 495 BP.
AC xxxxxx
DE Sequence 6, Application US/08222719
CC Sequence 6, Application US/08222719
CC Patent No. 5846711
CC GENERAL INFORMATION:
CC APPLICANT: David D. Moore
CC APPLICANT: Jae Moon Lee
CC TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
CC TITLE OF INVENTION: INTERACTING POLYPEPTIDES AND
CC TITLE OF INVENTION: RELATED MOLECULES AND METHODS
CC NUMBER OF SEQUENCES: 31
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.

CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: WordPerfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/222,719
CC FILING DATE: 04-April-1994
CC CLASSIFICATION: 436
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/969,136
CC FILING DATE: 30-October-1992
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Paul T. Clark
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 00786/229001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
SQ SEQUENCE 495 BP: 152 A; 130 C; 98 G; 114 T; 1 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 410 TTCTCAGAA 418
CP 33 ttctcagaa 25
RESULT 86
ID US-08-470-925-6 STANDARD; DNA; UNC; 495 BP.
AC xxxxxx
DE Sequence 6, Application US/08470925
CC Sequence 6, Application US/08470925
CC Patent No. 5866686
CC GENERAL INFORMATION:
CC APPLICANT: David D. Moore
CC APPLICANT: Jae Moon Lee
CC TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
CC TITLE OF INVENTION: INTERACTING POLYPEPTIDES AND
CC TITLE OF INVENTION: RELATED MOLECULES AND METHODS
CC NUMBER OF SEQUENCES: 31
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: WordPerfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/470,925
CC FILING DATE: 06-June-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/222,719
CC FILING DATE: 04-April-1994
CC CLASSIFICATION: 435

CC APPLICATION NUMBER: 07/969,136
CC FILING DATE: October 30, 1992
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Paul T. Clark
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 00786/229003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELETYPE: 200154
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC SEQUENCE 495 BP; 152 A; 130 C; 98 G; 114 T; 1 OTHER.
SQ
Query Match 100.0%; Score 9; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 410 TTCTCAGAA 418
|||||
CP 33 ttctcagaa 25
RESULT 87
PCT-US92-11353-13 STANDARD; DNA; UNC; 521 BP.
ID xxxxxx
AC
DE Sequence 13, Application PC/TUS9211353
CC Sequence 13, Application PC/TUS9211353
CC GENERAL INFORMATION:
CC APPLICANT: O'Hara, Patrick J
CC APPLICANT: Grant, Francis J
CC APPLICANT: Shepard, Paul O
CC TITLE OF INVENTION: NOVEL HUMAN TRANSGLUTAMINASES
CC NUMBER OF SEQUENCES: 22
CC FILING DATE: 19921230
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend
CC STREET: One Market Plaza, Stuart Street Tower
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105-1492
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/11353
CC FILING DATE: 19921230
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/816,284
CC FILING DATE: 31-DEC-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parmelee, Steve W
CC REGISTRATION NUMBER: 31-990
CC REFERENCE/DOCKET NUMBER: 13952-13-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 206-623-6793
CC TELEFAX: 206-467-9600
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 521 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC

CC CLONE: PTG562
SQ SEQUENCE 521 BP; 136 A; 125 C; 155 G; 105 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 521;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 349 TTCTCAGAA 357
|||||
CP 33 ttctcagaa 25
RESULT 88
US-07-998-973A-13 STANDARD; DNA; UNC; 521 BP.
ID xxxxxx
AC
DE Sequence 13, Application US/07998973A
CC Sequence 13, Application US/07998973A
CC Patent No. 5514579
CC GENERAL INFORMATION:
CC APPLICANT: O'Hara, Patrick J
CC APPLICANT: Grant, Francis J
CC APPLICANT: Shepard, Paul O
CC TITLE OF INVENTION: NOVEL HUMAN TRANSGLUTAMINASES
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend
CC STREET: One Market Plaza, Stuart Street Tower
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94105-1492
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/998,973A
CC FILING DATE: 19921230
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/816,284
CC FILING DATE: 31-DEC-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Parmelee, Steve W
CC REGISTRATION NUMBER: 31-990
CC REFERENCE/DOCKET NUMBER: 13952-13-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 206-623-6793
CC TELEFAX: 206-467-9600
CC INFORMATION FOR SEQ ID NO: 13:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 521 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: PTG562
SQ SEQUENCE 521 BP; 136 A; 125 C; 155 G; 105 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 521;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 349 TTCTCAGAA 357
|||||
CP 33 ttctcagaa 25
RESULT 89
US-08-171-385-4 STANDARD; DNA; UNC; 651 BP.
ID xxxxxx
AC

DE Sequence 4, Application US/08171385
CC Sequence 4, Application US/08171385
CC Patent No. 5527884
CC GENERAL INFORMATION:
CC APPLICANT: Mary E. Russell
CC APPLICANT: Ulrike Utans
CC TITLE OF INVENTION: Mediators of Chronic Allograft
CC TITLE OF INVENTION: Rejection
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: WordPerfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/171.385
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fraser, Janis K.
CC REGISTRATION NUMBER: 34,819
CC REFERENCE/DOCKET NUMBER: 05433/006001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 651
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
SQ SEQUENCE 651 BP; 218 A; 135 C; 168 G; 130 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 651;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 387 TTCACGAA 395
|||||
Cc 33 ttctcagaa 25

RESULT 90
ID US-08-322-742-16 STANDARD; DNA; UNC; 662 BP.
AC xxxxxx
DE Sequence 16, Application US/08322742
CC Sequence 16, Application US/08322742
CC Patent No. 5688641
CC GENERAL INFORMATION:
CC APPLICANT: Sager, Ruth
CC TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
CC SOFTWARE: WordPerfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/322,742
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,823
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: 07/844,296
CC FILING DATE: February 28, 1992
CC APPLICATION NUMBER: 07/552,216
CC FILING DATE: February 28, 1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fraser, Janis K.
CC REGISTRATION NUMBER: 34,819
CC REFERENCE/DOCKET NUMBER: 00530/048003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 662
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
SQ SEQUENCE 662 BP; 179 A; 158 C; 195 G; 130 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 460 TTCACGAA 468
|||||
Cc 33 ttctcagaa 25

RESULT 91
ID PCT-US94-01149-3 STANDARD; DNA; UNC; 666 BP.
AC xxxxxx
DE Sequence 3, Application PC/TUS9401149
CC Sequence 3, Application PC/TUS9401149
CC GENERAL INFORMATION:
CC APPLICANT: Shatzman, Allan
CC APPLICANT: Scott, Miller
CC APPLICANT: Dillon, Susan B.
CC APPLICANT: Kane, James
CC TITLE OF INVENTION: Vaccinal Polypeptides
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SmithKline Beecham Corporation - Corporate
CC ADDRESSEE: Patents
CC STREET: U.S. Mailcode UM2220 - 709 Swedeland Road
CC CITY: King Of Prussia
CC STATE: Pennsylvania
CC COUNTRY: USA
CC ZIP: 19406-2799
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01149
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 149,150
CC FILING DATE: 05-NOV-1993
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 013,415
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 108,914
CC FILING DATE: 18-AUG-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 837,773
CC FILING DATE: 18-FEB-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 751,896
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 387,200
CC FILING DATE: 28-JUL-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 238,801
CC FILING DATE: 02-NOV-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 645,732
CC FILING DATE: 30-AUG-1984
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Baumeister, Kirk
CC REGISTRATION NUMBER: 33,833
CC REFERENCE/DOCKET NUMBER: P50134 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-270-5096
CC TELEFAX: 215-270-5090
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 666 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..663
CC SEQUENCE 666 BP: 224 A; 118 C; 166 G; 158 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 208 TTCTCAGAA 216
Cp 33 ttctcagaa 25
RESULT 92
ID PCT-US94-01149-1 STANDARD; DNA; UNC; 666 BP.
AC xxxxxx
DE Sequence 1, Application PC/TUS9401149
CC Sequence 1, Application PC/TUS9401149
CC GENERAL INFORMATION:
CC APPLICANT: Shatzman, Allan
CC APPLICANT: Scott, Miller
CC APPLICANT: Dillon, Susan B.
CC APPLICANT: Kane, James
CC TITLE OF INVENTION: Vaccinal Polypeptides
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Smithkline Beecham Corporation - Corporate
CC STREET: U.S. Mailcode UM2220 - 709 Swedeland Road
CC CITY: King of Prussia
CC STATE: Pennsylvania
CC COUNTRY: USA
CC ZIP: 19406-2799
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01149
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 149,150
CC FILING DATE: 05-NOV-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 013,415
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 108,914
CC FILING DATE: 18-AUG-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 837,773
CC FILING DATE: 18-FEB-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 751,896
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 387,200
CC FILING DATE: 28-JUL-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 238,801
CC FILING DATE: 02-NOV-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 645,732
CC FILING DATE: 30-AUG-1984
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Baumeister, Kirk
CC REGISTRATION NUMBER: 33,833
CC REFERENCE/DOCKET NUMBER: P50134 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-270-5096
CC TELEFAX: 215-270-5090
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 666 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..663
CC SEQUENCE 666 BP: 223 A; 117 C; 167 G; 159 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 208 TTCTCAGAA 216
Cp 33 ttctcagaa 25
RESULT 93
ID PCT-US94-01149-7 STANDARD; DNA; UNC; 670 BP.
AC xxxxxx
DE Sequence 7, Application PC/TUS9401149
CC Sequence 7, Application PC/TUS9401149
CC GENERAL INFORMATION:
CC APPLICANT: Shatzman, Allan
CC APPLICANT: Scott, Miller
CC APPLICANT: Dillon, Susan B.
CC APPLICANT: Kane, James
CC TITLE OF INVENTION: Vaccinal Polypeptides
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Smithkline Beecham Corporation - Corporate
CC STREET: Patents

CC STREET: U.S. Mailcode UW2220 - 709 Swedeland Road
CC CITY: King of Prussia
CC STATE: Pennsylvania
CC COUNTRY: USA
CC ZIP: 19406-2799
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01149
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 149,150
CC FILING DATE: 05-NOV-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 013,415
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 108,914
CC FILING DATE: 18-AUG-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 837,773
CC FILING DATE: 18-FEB-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 751,896
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 387,200
CC FILING DATE: 28-JUL-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 238,801
CC FILING DATE: 02-NOV-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 645,732
CC FILING DATE: 30-AUG-1984
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Baumeister, Kirk
CC REGISTRATION NUMBER: 33,833
CC REFERENCE/DOCKET NUMBER: P50134 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-270-5096
CC TELEFAX: 215-270-5090
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 670 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..670
CC SEQUENCE 670 BP: 117 C; 166 G; 158 T; 6 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 670;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 TTCTCAGAA 216
|||||
Cp 33 ttctcagaa 25

RESULT 94
ID US-08-139-937-9 STANDARD; DNA; UNC; 688 BP.
AC xxxxxx
DE Sequence 9, Application US/08139937
CC Sequence 9, Application US/08139937
CC Patent No. 5821070

CC GENERAL INFORMATION:
CC APPLICANT: LEE, WEN-HWA
CC APPLICANT: SHAN, BEI
CC TITLE OF INVENTION: CELLULAR GENES ENCODING
CC TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: CAMPBELL AND FLORES
CC STREET: 4370 LA JOLLA VILLAGE DRIVE
CC CITY: SAN DIEGO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 92122
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/139,937
CC FILING DATE: 20-OCT-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/979,156
CC FILING DATE: 20-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CAMPBELL, CATRYN
CC REGISTRATION NUMBER: 31,815
CC REFERENCE/DOCKET NUMBER: P-CI 9370
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-535-9001
CC TELEFAX: 619-535-8949
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 688 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 688 BP: 253 A; 116 C; 149 G; 170 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 688;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 292 TTCTCAGAA 300
|||||
Cp 33 ttctcagaa 25

RESULT 95
ID PCT-US93-11310-9 STANDARD; DNA; UNC; 688 BP.
AC xxxxxx
DE Sequence 9, Application PC/TUS9311310
CC Sequence 9, Application PC/TUS9311310
CC GENERAL INFORMATION:
CC APPLICANT: BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM
CC TITLE OF INVENTION: CELLULAR GENES ENCODING
CC TITLE OF INVENTION: RETINOBLASTOMA-ASSOCIATED PROTEINS
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: CAMPBELL AND FLORES
CC STREET: 4370 LA JOLLA VILLAGE DRIVE
CC CITY: SAN DIEGO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 92122
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/11310
CC FILING DATE: 19-NOV-1993
CC
CC - CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CAMPBELL, CATHERYN
CC REGISTRATION NUMBER: 31,815
CC REFERENCE/DOCKET NUMBER: FP-CJ 9790
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619-535-9001
CC TELEFAX: 619-535-8949
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 688 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 688 BP; 253 A; 116 C; 149 G; 170 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 688;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 292 TTCTCAGAA 300
CP 33 ttctcagaa 25

RESULT 96
ID PCT-US96-03940-2 STANDARD; DNA; UNC; 756 BP.
AC xxxxxx
DT Sequence 2, Application PC/TUS9603940
DE Sequence 2, Application PC/TUS9603940
CC GENERAL INFORMATION:
CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC TITLE OF INVENTION: RESR Protein and DNA
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechert Price & Rhoads
CC STREET: P. O. Box 5218
CC CITY: Princeton
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 08543-5218
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
CC COMPUTER: IBM-compatible
CC OPERATING SYSTEM: DOS 5.0
CC SOFTWARE: Wordperfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03940
CC FILING DATE: March 23, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Allen Bloom
CC REGISTRATION NUMBER: 29,135
CC REFERENCE/DOCKET NUMBER: 317743-101
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214
CC TELEFAX: (609) 520-3259
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 756 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: no
CC ANTI-SENSE: no
CC ORIGINAL SOURCE:
CC ORGANISM: Human

CC CELL LINE: HeLa
CC IMMEDIATE SOURCE:
CC LIBRARY: CDNA
CC PUBLICATION INFORMATION:
CC AUTHORS: Chong, Jayhong A., Tapia-Ramirez Jos, Toledo-
CC AUTHORS: Aral, Juan, Zhang, Yingsong, Boutros, Michael C., Altschuler,
CC AUTHORS: Yelena M., Frohman, Michael A., Krainer, Susan D., Mandel, Gail
CC TITLE: REST: A Mammalian Silencer Protein that Restricts
CC JOURNAL: Cell
CC VOLUME: 80
CC ISSUE:
CC PAGES:
CC DATE: March 24, 1995
CC RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 756
SQ SEQUENCE 756 BP; 259 A; 164 C; 153 G; 180 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 459 TTCTCAGAA 467
CP 33 ttctcagaa 25

RESULT 97
ID US-08-216-233C-1 STANDARD; DNA; UNC; 776 BP.
AC xxxxxx
DT Sequence 1, Application US/08216233C
DE Sequence 1, Application US/08216233C
CC Patent No. 5506105
CC GENERAL INFORMATION:
CC APPLICANT: HADDOCK, PAUL V.
CC TITLE OF INVENTION: IN SITU ASSAY OF AMPLIFIED INTRACELLULAR
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DADE INTERNATIONAL INC.
CC STREET: 1717 DEERFIELD ROAD
CC CITY: DEERFIELD
CC STATE: ILLINOIS
CC COUNTRY: USA
CC ZIP: 60015
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/216,233C
CC FILING DATE: 22-MAR-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/808,456
CC FILING DATE: 10-DEC-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: TYMESON, CYNTHIA G.
CC REGISTRATION NUMBER: 34,745
CC REFERENCE/DOCKET NUMBER: BA-4203
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (305) 222-6423
CC TELEFAX: (305) 222-6686
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 776 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 776 BP; 260 A; 138 C; 170 G; 208 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 776;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 761 TTCTCAGAA 769
|||||
Cp 33 ttctcagaa 25

RESULT 98
ID US-08-117-083-7 STANDARD; DNA; UNC; 790 BP.
AC xxxxxx
DE Sequence 7, Application US/08117083
CC Sequence 7, Application US/08117083
CC Patent No. 5719054
CC GENERAL INFORMATION:
CC APPLICANT: Bournell, Michael E.
CC APPLICANT: Ingils, Stephen C.
CC APPLICANT: Munro, Alan J.
CC TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
CC NUMBER OF SEQUENCES: 70
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Walter H. Dreyer
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/117,083
CC FILING DATE: 10-SEP-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dreyer, Walter H.
CC REGISTRATION NUMBER: 24,190
CC REFERENCE/DOCKET NUMBER: A-58783
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-781-1989
CC TELEFAX: 415-398-3249
CC FAX: 910 277299
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 790 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 790 BP: 262 A; 143 C; 175 G; 210 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 766 TTCTCAGAA 774
|||||
Cp 33 ttctcagaa 25

RESULT 99
ID US-08-565-386-5 STANDARD; DNA; UNC; 792 BP.
AC xxxxxx
DE Sequence 5, Application US/08565386
CC Sequence 5, Application US/08565386
CC Patent No. 5741697
CC TELECOMMUNICATION INFORMATION:
CC GENERAL INFORMATION:

CC APPLICANT: Bayoll, Patrick M.
CC APPLICANT: Hsia, Ru-ching
CC TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
CC STREET: Clinton Square, P.O. Box 1051
CC CITY: Roches
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/565,386
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Timian, Susan J.
CC REGISTRATION NUMBER: 34,103
CC REFERENCE/DOCKET NUMBER: 176/60040
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 716-263-1636

Query Match 100.0%; Score 9; DB 2; Length 792;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 336 TTCTCAGAA 344
|||||
Cp 25 ttctcagaa 33

RESULT 100
ID US-08-565-386-5 STANDARD; DNA; UNC; 792 BP.
AC xxxxxx
DE Sequence 5, Application US/08565386
CC Sequence 5, Application US/08565386
CC Patent No. 5741697
CC GENERAL INFORMATION:
CC APPLICANT: Bayoll, Patrick M.
CC APPLICANT: Hsia, Ru-ching
CC TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
CC STREET: Clinton Square, P.O. Box 1051
CC CITY: Roches
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/565,386
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Timian, Susan J.
CC REGISTRATION NUMBER: 34,103
CC REFERENCE/DOCKET NUMBER: 176/60040
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 716-263-1636

CC TELEFAX: 716-263-1600
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 792 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: circular
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 792 BP: 226 A; 133 C; 184 G; 249 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 792;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 84 TTCTCAGAA 92
33 ttctcagaa 25

RESULT 101
ID US-08-967-101-119 STANDARD; DNA; UNC; 796 BP.
AC xxxxxx

DE Sequence 119, Application US/08967101
CC Sequence 119, Application US/08967101
CC Patent No. 5840540

GENERAL INFORMATION:
CC APPLICANT: ST. GEORGE-HYSLOP, PETER H
CC APPLICANT: ROMENS, JOHANNA M
CC APPLICANT: FRASER, PAUL E
CC TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
CC TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
CC NUMBER OF SEQUENCES: 183

CORRESPONDENCE ADDRESS:
CC ADDRESSEE: TESTA, HURWITZ & THIBEAULT
CC STREET: High Street Tower - 125 High Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110

COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/967,101

FILING DATE: 10-NOV-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/592,541
CC FILING DATE:

ATTORNEY/AGENT INFORMATION:
CC NAME: Pitcher, Edmund R.
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 248-7000
CC TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 119:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 796 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 796 BP: 198 A; 134 C; 264 G; 182 T; 18 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 796;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 450 TTCTCAGAA 458
33 ttctcagaa 25

RESULT 102
ID PCT-US92-06617A-2 STANDARD; DNA; UNC; 796 BP.
AC xxxxxx

DE Sequence 2, Application PC/TUS9206617A
CC Sequence 2, Application PC/TUS9206617A
CC GENERAL INFORMATION:

CC APPLICANT: Parodos, Kyriaki
CC APPLICANT: McCarthy, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC TITLE OF INVENTION: Shigella
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Amoco Corporation
CC STREET: 200 East Randolph Drive, P.O. Box 87703
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: U.S.A.
CC ZIP: 60680

COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/06617A
CC FILING DATE: 19920728

PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Galloway, Norval B.
CC REGISTRATION NUMBER: 33,595
CC REFERENCE/DOCKET NUMBER: GTR90-04 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312-856-7180
CC TELEFAX: 312-856-4972

INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 796 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 796 BP: 246 A; 165 C; 187 G; 198 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 796;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 TTCTCAGAA 41
33 ttctcagaa 25

RESULT 103
ID US-08-375-241-2 STANDARD; DNA; UNC; 796 BP.
AC xxxxxx

DE Sequence 2, Application US/08375241
CC Sequence 2, Application US/08375241
CC Patent No. 5648481

GENERAL INFORMATION:
CC APPLICANT: Parodos, Kyriaki
CC APPLICANT: McCarthy, Janice
CC TITLE OF INVENTION: Nucleic Acid Probes for the Detection of
CC TITLE OF INVENTION: Shigella
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington

CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02113
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/375,241
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/738,800
CC FILING DATE: 31-JUL-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: GTR90-04
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 796 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC SEQUENCE 796 BP: 165 C; 187 G; 198 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 796;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 33 TTCGACAA 41
33 ttctcagaa 25
CP
RESULT 104
ID US-08-407-544-1 STANDARD; DNA; UNC; 798 BP.
AC xxxxxx
DT Sequence 1, Application US/08407544
DE Sequence 1, Application US/08407544
CC Patent No. 5741657
CC GENERAL INFORMATION:
CC APPLICANT: Tsien, Roger Y.
CC APPLICANT: Zlokarnik, Gregor
CC TITLE OF INVENTION: FLUOROGENIC SUBSTRATES FOR
CC TITLE OF INVENTION: BETA-LACTAMASE AND METHODS FOR THE PREPARATION AND USE
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: C/O Robbins, Berliner & Carson
CC STREET: 210 N. Figueroa Street, Suite 500
CC CITY: Los Angeles
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 90012
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/407,544
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Spitals, John P.
CC REGISTRATION NUMBER: 29,215
CC REFERENCE/DOCKET NUMBER: 1279-197

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 977-1001
CC TELEFAX: (213) 977-1003
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 798 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 7..795
SO SEQUENCE 798 BP: 217 A; 185 C; 213 G; 183 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 798;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 222 TTCGACAA 230
33 ttctcagaa 25
CP
RESULT 105
ID US-08-967-101-118 STANDARD; DNA; UNC; 819 BP.
AC xxxxxx
DT Sequence 118, Application US/08867101
DE Sequence 118, Application US/08867101
CC Patent No. 5840540
CC GENERAL INFORMATION:
CC APPLICANT: ST. GEORGE-HYSLOP, PETER H
CC APPLICANT: ROMMENS, JOHANNA M
CC APPLICANT: FRASER, PAUL E
CC TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
CC TITLE OF INVENTION: TO ALZHEIMER'S DISEASE
CC NUMBER OF SEQUENCES: 183
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: TESTA, HURWITZ & THIBEAULT
CC STREET: High Street Tower - 125 High Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/967,101
CC FILING DATE: 10-NOV-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/592,541
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Pitcher, Edmund R.
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 248-7000
CC TELEFAX: (617) 248-7100
CC INFORMATION FOR SEQ ID NO: 118:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 819 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SO SEQUENCE 819 BP: 185 A; 232 C; 162 G; 208 T; 32 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 819;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 427 TTCGTGAGAA 435

OY 25 ttctcgagaa 33

RESULT 106
ID US-08-485-284A-4 STANDARD; DNA; UNC; 850 BP.

AC xxxxxx

DE Sequence 4, Application US/08485284A

CC Sequence 4, Application US/08485284A

CC Patent No. 5750372

CC GENERAL INFORMATION:

CC APPLICANT: SAKAI, YASUYOSHI

CC APPLICANT: TANI, YOSHIKI

CC APPLICANT: SHIBANO, YUJI

CC APPLICANT: KONO, HIROMI

CC APPLICANT: HATAKAKA, HARUYO

CC TITLE OF INVENTION: NOVEL VECTOR HAVING PROMOTER THAT IS

CC NUMBER OF SEQUENCES: 18

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.

CC STREET: 1100 New York Avenue, N.W.

CC CITY: Washington

CC STATE: D. C.

CC COUNTRY: U.S.A.

CC ZIP: 20005-3918

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/485,284A

CC FILING DATE: 07-JUN-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: JP 43361/1992

CC FILING DATE: 28-FEB-1992

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/025,416

CC FILING DATE: 01-MAR-1993

CC ATTORNEY/AGENT INFORMATION:

CC NAME: WHITE JR., PAUL E.

CC REGISTRATION NUMBER: 32,011

CC TELEPHONE: (202) 861-3000

CC TELEFAX: (202) 822-0944

CC TELEFAX: 6714627 CUSH

CC INFORMATION FOR SEQ ID NO: 4:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 850 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double

CC TOPOLOGY: unknown

CC MOLECULE TYPE: genomic DNA

CC SEQUENCE 850 BP; 351 A; 83 C; 121 G; 295 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 850;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 337 TTCGTGAGAA 345

OY 25 ttctcgagaa 33

RESULT 107
ID US-08-247-809A-11 STANDARD; DNA; UNC; 850 BP.

AC xxxxxx

DE Sequence 11, Application US/08247809A

CC Sequence 11, Application US/08247809A

CC Patent No. 5569823

CC GENERAL INFORMATION:

CC APPLICANT: Peter H. Schreier; Klaus Stenzel; Gunter Adam;

CC APPLICANT: Edgar Maiss

CC TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS

CC NUMBER OF SEQUENCES: 18

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: SPRUNG HORN KRAMER & WOODS

CC STREET: 660 White Plains Road

CC CITY: Tarrytown

CC STATE: New York

CC COUNTRY: U.S.A.

CC ZIP: 10591-5144

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB

CC OPERATING SYSTEM: DOS

CC SOFTWARE: WordPerfect 5.1

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/247,809A

CC FILING DATE: May 23, 1994

CC CLASSIFICATION: 800

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: P 43 178 45.6 (Germany)

CC FILING DATE: May 28, 1993

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Kurt G. Briscoe

CC REGISTRATION NUMBER: 33,141

CC TELEPHONE: (914) 332-1700

CC TELEFAX: (914) 332-1844

CC INFORMATION FOR SEQ ID NO: 11:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 850 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 2..799

CC SEQUENCE 850 BP; 291 A; 145 C; 180 G; 234 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 850;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 355 TTCGTGAGAA 363

OY 25 ttctcgagaa 33

RESULT 108
ID US-08-346-333-15 STANDARD; DNA; UNC; 858 BP.

AC xxxxxx

DE Sequence 15, Application US/08346333

CC Sequence 15, Application US/08346333

CC Patent No. 5677153

CC GENERAL INFORMATION:

CC APPLICANT: Botstein, David

CC APPLICANT: Palzkill, Timothy

CC TITLE OF INVENTION: Methods for modifying DNA and for

CC TITLE OF INVENTION: detecting effects of such modification on interaction of

CC NUMBER OF SEQUENCES: 86

```

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Richard F. Trecartin
CC STREET: 4 Embarcadero Center, Suite 3400
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patemlin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/346.333
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/039.501
CC FILING DATE:
CC APPLICATION NUMBER: US 07/602.158
CC FILING DATE: 22-OCT-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Trecartin, Richard F.
CC REGISTRATION NUMBER: 31,801
CC REFERENCE/DOCKET NUMBER: A-53469/RFT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC INFORMATION FOR SEQ ID NO: 15:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 858 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..858
SQ SEQUENCE 858 BP; 221 A; 205 C; 220 G; 212 T; 0 OTHER.
CC
CC Query Match 100.0%; Score 9; DB 2; Length 858;
CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dn 285 TTCACGAA 293
Cp 33 ttctcgaa 25

```

```

CC CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: PCT/US91/07506
CC CC FILING DATE: 19911021
CC CC CLASSIFICATION: 435
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Treccartin, Richard F.
CC CC REGISTRATION NUMBER: 31,801
CC CC REFERENCE/DOCKET NUMBER: FP-53469-PC/RFT
CC CC TELECOMMUNICATION INFORMATION:
CC CC TELEPHONE: (415) 781-1989
CC CC TELEFAX: (415) 398-3249
CC CC INFORMATION FOR SEQ ID NO: 15:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 858 base pairs
CC CC TYPE: NUCLEIC ACID
CC CC STRANDEDNESS: single
CC CC TOPOLOGY: linear
CC CC MOLECULE TYPE: CDNA
CC CC FEATURE:
CC CC NAME/KEY: CDS
CC CC LOCATION: 1..858
SQ SEQUENCE 858 BP: 221 A; 205 C; 220 G; 212 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 858;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 TTCCTCAGAA 293
CP 33 ttctcagaa 25

RESULT 110
ID US-08-406-849-1 STANDARD; DNA; UNC; 861 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08406849
CC Sequence 1, Application US/08406849
CC Patent No. 5849547
CC GENERAL INFORMATION:
CC APPLICANT: CLEZIZAT, Philippe
CC APPLICANT: GUILLOU-BONNICI, Francois
CC APPLICANT: MALLET, Francois
CC APPLICANT: LEVASSEUR, Pierre
CC TITLE OF INVENTION: METHOD FOR NUCLEIC ACID AMPLIFICATION BY
CC TITLE OF INVENTION: TRANSCRIPTION USING DISPLACEMENT, AND REAGENTS AND KIT
CC TITLE OF INVENTION: THEREFOR
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OLIVIER & BERRIDGE
CC STREET: 700 South Washington Street, Suite 300
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22314
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/406,849
CC FILING DATE: April 24, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berridge, William P.
CC REGISTRATION NUMBER: 30,024
CC REFERENCE/DOCKET NUMBER: WPB 36251
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-836-6400
CC TELEFAX: 703-836-2787
CC INFORMATION FOR SEQ ID NO: 1:

```


CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 861 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "DNA"
SQ SEQUENCE 861 BP; 223 A; 203 C; 222 G; 213 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 861;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 285 TTCTCAGAA 293
Cp 33 ttctcagaa 25

RESULT 111
ID US-08-650-578-1 STANDARD; DNA; UNC; 871 BP.
AC xxxxxx
DT Sequence 1, Application US/08650578
CC Sequence 1, Application US/08650578
CC Patent No. 5811284
CC GENERAL INFORMATION:
CC APPLICANT: Chang, Chiwen
CC APPLICANT: Aramburu Beltiran, Jose
CC APPLICANT: Lopez-Botet, Miguel
CC APPLICANT: Phillips Jr., Joseph H.
CC APPLICANT: Ianier, Lewis L.
CC TITLE OF INVENTION: Purified Mammalian NK Antigens and
CC TITLE OF INVENTION: Related Reagents
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: DNAX Research Institute
CC STREET: 901 California Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304-1104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/650,578
CC FILING DATE:
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/175,339
CC FILING DATE: 29-DEC-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ching, Edwin P.
CC REGISTRATION NUMBER: 34,090
CC REFERENCE/DOCKET NUMBER: DX0391
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-852-9196
CC TELEFAX: 415-496-1200
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 871 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 122..661
SQ SEQUENCE 871 BP; 263 A; 168 C; 168 G; 272 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 871;

Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 391 TTCTCAGAA 399
Cp 33 ttctcagaa 25

RESULT 112
ID 5258283-1 STANDARD; DNA; UNC; 948 BP.
AC xxxxxx
DT 01-JAN-1900
DE Patent No. 5258283.
CC Patent No. 5258283.
CC APPLICANT: FRAZIER, MARVIN E.; MALLAVIA, LOUIS P.; SAMUEL,
CC JAMES E.; BACA, OSWALD G.
CC TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF COXIELLA
CC BURNETII IN BIOLOGICAL FLUIDS
CC NUMBER OF SEQUENCES: 17
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/425,856
CC FILING DATE: 23-OCT-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 927,779
CC FILING DATE: 05-NOV-1986
CC APPLICATION NUMBER: 795,207
CC FILING DATE: 05-NOV-1985
CC SEQ ID NO: 1:
CC LENGTH: 875
SQ Sequence 948 BP; 259 A; 161 C; 172 G; 283 T; 73 other;

Query Match 100.0%; Score 9; DB 5; Length 875;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 364 TTCTCAGAA 372
Cp 33 ttctcagaa 25

RESULT 113
ID US-08-158-682A-3 STANDARD; DNA; UNC; 879 BP.
AC xxxxxx
DT Sequence 3, Application US/08158682A
CC Sequence 3, Application US/08158682A
CC Patent No. 5434058
CC GENERAL INFORMATION:
CC APPLICANT: Davidson, Nicholas O.
CC TITLE OF INVENTION: Apolipoprotein B RNA Editing Protein;
CC TITLE OF INVENTION: Composition and Method
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: ARNOLD, WHITE & DURKEE
CC STREET: 321 No. 5434058th Clark Street, Suite 800
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60610
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/158,682A
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Cooley, Ronald B.
CC REGISTRATION NUMBER: 27,187
CC REFERENCE/DOCKET NUMBER: ARCD:085
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (312) 744-0090
CC TELEFAX: (312) 245-4961
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 879 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 23..730
SQ SEQUENCE 879 BP; 254 A; 207 C; 190 G; 228 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 879;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 TTCGAGAA 36
|||
QY 25 ttctgagaa 33

RESULT 114
ID US-08-801-740-4 STANDARD; DNA; UNC; 892 BP.
AC xxxxxx
DE Sequence 4, Application US/08801740
CC Sequence 4, Application US/08801740
CC Patent No. 5869639
CC GENERAL INFORMATION:
CC APPLICANT: Goli, Surya K.
CC TITLE OF INVENTION: NOVEL P24 VESICLE PROTEINS
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/801,740
CC FILING DATE: Herewith
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: P-0189 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 892 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: Consensus
SQ SEQUENCE 892 BP; 268 A; 192 C; 210 G; 221 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 892;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 432 TTCGAGAA 440
|||
QY 25 ttctgagaa 33

RESULT 115
ID PCT-US94-01149-9 STANDARD; DNA; UNC; 918 BP.
AC xxxxxx
DE Sequence 9, Application PC/TUS9401149
CC Sequence 9, Application PC/TUS9401149
CC GENERAL INFORMATION:
CC APPLICANT: Shatzman, Allan
CC APPLICANT: Scott, Miller
CC APPLICANT: Dillon, Susan B.
CC APPLICANT: Kane, James
CC TITLE OF INVENTION: Vaccinal Polypeptides
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SmithKline Beecham Corporation - Corporate
CC ADDRESSEE: Patents
CC STREET: U.S. Mailcode UW2220 - 709 Swedeland Road
CC CITY: King of Prussia
CC STATE: Pennsylvania
CC COUNTRY: USA
CC ZIP: 19406-2799
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01149
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 149,150
CC FILING DATE: 05-NOV-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 013,415
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 108,914
CC FILING DATE: 18-AUG-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 837,773
CC FILING DATE: 18-FEB-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 751,896
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 387,200
CC FILING DATE: 28-JUL-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 236,801
CC FILING DATE: 02-NOV-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 645,732
CC FILING DATE: 30-AUG-1984
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Baumeister, Kirk
CC REGISTRATION NUMBER: 33,833
CC REFERENCE/DOCKET NUMBER: P50134 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-270-5096
CC TELEFAX: 215-270-5090
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 918 base pairs

CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..918
SQ SEQUENCE 918 BP: 295 A; 176 C; 234 G; 213 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 918;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 463 TTCTCAGAA 471
|||||
CP 33 ttctcagaa 25

RESULT 116
ID PCT-US94-01149-58 STANDARD; DNA; UNC; 918 BP.
AC xxxxxx

DE Sequence 58, Application PC/TUS9401149
CC Sequence 58, Application PC/TUS9401149
CC GENERAL INFORMATION:
CC APPLICANT: Shatzman, Allan
CC APPLICANT: Scott, Miller
CC APPLICANT: Dillon, Susan B.
CC APPLICANT: Kane, James
CC TITLE OF INVENTION: Vaccinal Polypeptides
CC NUMBER OF SEQUENCES: 72
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SmithKline Beecham Corporation - Corporate
CC ADDRESSEE: Patents
CC STREET: U.S. Mailcode UW2220 - 709 Swedeland Road
CC CITY: King of Prussia
CC STATE: Pennsylvania
CC COUNTRY: USA
CC ZIP: 19406-2799

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/01149
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 149,150
CC FILING DATE: 05-NOV-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 013,415
CC FILING DATE: 01-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 108,914
CC FILING DATE: 18-AUG-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 837,773
CC FILING DATE: 18-FEB-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 751,896
CC FILING DATE: 30-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 387,200
CC FILING DATE: 28-JUL-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 238,801
CC FILING DATE: 02-NOV-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 645,732
CC FILING DATE: 30-AUG-1984
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Baumeister, Kirk
CC REGISTRATION NUMBER: 33,833
CC REFERENCE/DOCKET NUMBER: P50134 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 215-270-5096
CC TELEFAX: 215-270-5090
CC INFORMATION FOR SEQ ID NO: 58:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 918 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 918 BP: 292 A; 179 C; 231 G; 216 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 918;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 463 TTCTCAGAA 471
|||||
CP 33 ttctcagaa 25

RESULT 117
ID PCT-US96-03940-5 STANDARD; DNA; UNC; 928 BP.
AC xxxxxx

DE Sequence 5, Application PC/TUS9603940
CC Sequence 5, Application PC/TUS9603940
CC GENERAL INFORMATION:
CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC TITLE OF INVENTION: REST Protein and DNA
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechert Price & Rhoads
CC STREET: P.O. Box 5218
CC CITY: Princeton
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 08543-5218

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
CC COMPUTER: IBM-compatible
CC OPERATING SYSTEM: DOS 5.0
CC SOFTWARE: WordPerfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03940
CC FILING DATE: March 23, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Allen Bloom
CC REGISTRATION NUMBER: 29,135
CC REFERENCE/DOCKET NUMBER: 317743-101
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214
CC TELEFAX: (609) 520-3259
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 928 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: no
CC ANTI-SENSE: no
CC ORIGINAL SOURCE:
CC ORGANISM: Human
CC CELL LINE: HeLa
CC IMMEDIATE SOURCE:
CC LIBRARY: CDNA
CC PUBLICATION INFORMATION:
CC AUTHORS: Chong, Jayhong A., Tapia-Ramirez Jos, Toledo-

CC AUTHORS: Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
CC AUTHORS: Yelena M., Fromman, Michael A., Krainer, Susan D., Mandel, Gail
CC TITLE: REST: A Mammalian Silencer Protein that Restricts
CC TITLE: Sodium Channel Gene Expression to Neurons
CC JOURNAL: Cell
CC VOLUME: 80
CC ISSUE:
CC PAGES:
CC DATE: March 24, 1995
CC RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 928
SQ SEQUENCE 928 BP: 363 A; 178 C; 182 G; 205 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 928;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 245 TTCACGAA 253
Cp 33 ttctcagaa 25

RESULT 118
ID US-08-632-514C-9 STANDARD; DNA; UNC; 958 BP.
AC xxxxxx
DE Sequence 9, Application US/08632514C
CC Sequence 9, Application US/08632514C
CC Patent No. 5834234
CC GENERAL INFORMATION:
CC APPLICANT: GALLO, Gregory J.
CC TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BIK
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hale and Dorr
CC STREET: 1455 Pennsylvania Avenue, N.W.
CC CITY: Washington, D.C.
CC COUNTRY: U.S.A.
CC ZIP: 20004
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/632,514C
CC FILING DATE: 29-MAY-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WIXON, Henry N.
CC REGISTRATION NUMBER: 32,073
CC REFERENCE/DOCKET NUMBER: 104322.188
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)942-8459
CC TELEFAX: (202)942-8484
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 958 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 958 BP: 163 A; 314 C; 282 G; 199 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 958;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 806 TTCACGAA 814
Cp 33 ttctcagaa 25

RESULT 119

ID US-08-888-171-12 STANDARD; DNA; UNC; 969 BP.
AC xxxxxx
DE Sequence 12, Application US/08888171
CC Sequence 12, Application US/08888171
CC Patent No. 5851528
CC GENERAL INFORMATION:
CC APPLICANT: Joone-Long, Ko
CC APPLICANT: Higgins, Paul J.
CC APPLICANT: Yen, C. Grace
CC TITLE OF INVENTION: METHODS OF INHIBITING COMPLEMENT
CC TITLE OF INVENTION: ACTIVATION
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson, P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: US
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: Windows95
CC SOFTWARE: FastSeq for Windows version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/888,171
CC FILING DATE: 03-JUL-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/310,416
CC FILING DATE: 22-SEP-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Freeman, John W.
CC REGISTRATION NUMBER: 29,066
CC REFERENCE/DOCKET NUMBER: 06180/005002
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/542-507
CC TELEFAX: 617/542-890
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 969 BP: 308 A; 224 C; 198 G; 239 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 969;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 834 TTCACGAA 842
Cp 33 ttctcagaa 25

RESULT 120
ID US-08-310-416A-12 STANDARD; DNA; UNC; 969 BP.
AC xxxxxx
DE Sequence 12, Application US/08310416A
CC Sequence 12, Application US/08310416A
CC Patent No. 5679546
CC GENERAL INFORMATION:
CC APPLICANT: Joone-Long, Ko et al.
CC TITLE OF INVENTION: CHIMERIC PROTEINS WHICH BLOCK
CC TITLE OF INVENTION: COMPLEMENT ACTIVATION
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston

CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: MS-DOS (Version 5.0)
CC SOFTWARE: Wordperfect (Version 5.1)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/310.416A
CC FILING DATE: 22-SEP-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Paul T. Clark
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 06180/005001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
SQ SEQUENCE 969 BP: 308 A; 224 C; 198 G; 239 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 969;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 834 TTCCTCAGAA 842
CP 33 ttctcagaa 25
RESULT 121
ID US-08-243-543A-2 STANDARD; DNA; UNC; 972 BP.
AC xxxxxx
DE Sequence 2, Application US/08243543A
CC Sequence 2, Application US/08243543A
CC Patent No. 5614364
CC GENERAL INFORMATION:
CC APPLICANT: Tuggle, Christopher K
CC TITLE OF INVENTION: Genetic Marker for Improved Milk
CC TITLE OF INVENTION: Production Traits in Cattle
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
CC STREET: 801 Grand Avenue Suite 3200
CC CITY: Des Moines
CC STATE: Iowa
CC COUNTRY: United States
CC ZIP: 50309
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/243.543A
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Nebel, Heidi S.
CC REGISTRATION NUMBER: 37,719
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 515-288-3667
CC TELEFAX: 515-288-1338
CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 972 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 466..678
CC FEATURE:
CC NAME/KEY: allele
CC LOCATION: replace(630, "g")
CC FEATURE:
CC NAME/KEY: Intron
CC LOCATION: 1..465
CC FEATURE:
CC NAME/KEY: Intron
CC LOCATION: 679..972
SQ SEQUENCE 972 BP: 288 A; 190 C; 176 G; 315 T; 3 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 972;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 851 TTCCTCAGAA 859
CP 33 ttctcagaa 25
RESULT 122
ID US-08-592-126-57 STANDARD; DNA; UNC; 987 BP.
AC xxxxxx
DE Sequence 57, Application US/08592126
CC Sequence 57, Application US/08592126
CC Patent No. 5821091
CC GENERAL INFORMATION:
CC APPLICANT: Gregory Dolganov
CC TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
CC TITLE OF INVENTION: Polypeptides
CC NUMBER OF SEQUENCES: 151
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/592.126
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles R.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 4600-0111
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 57:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 987 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA

CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: TC2.seq
SQ SEQUENCE 987 BP; 267 A; 223 C; 223 G; 274 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 987;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 9 TTCGAGAA 17
25 ttctgagaa 33
OY
RESULT 123
ID US-07-965-274-3 STANDARD; DNA: UNC; 1000 BP.
AC xxxxxx
DE Sequence 3, Application US/07965274
CC Sequence 3, Application US/07965274
CC Patent No. 5501947
CC GENERAL INFORMATION:
CC APPLICANT: EMERY, VINCENT C.
CC APPLICANT: BAVIN, PEGGY J.
CC APPLICANT: WALKER, PATRICK
CC TITLE OF INVENTION: METHOD FOR THE DIAGNOSIS OF HUMAN
CC TITLE OF INVENTION: PAPILLOMA VIRUS TYPE 16 USING THE POLYMERASE CHAIN
CC NUMBER OF SEQUENCES: 3
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: REED & ROBINS
CC STREET: 635 BRYANT STREET
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: UNITED STATES OF AMERICA
CC ZIP: 94301
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/965,274
CC FILING DATE: 12-MAR-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/GB91/01212
CC FILING DATE: 19-JUL-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: GB 9015845.2
CC FILING DATE: 19-JUL-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: ROBINS, ROBERTA L.
CC REGISTRATION NUMBER: 33,208
CC REFERENCE/DOCKET NUMBER: 5150-0028
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 617-8999
CC TELEFAX: (415) 327-3231
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1000 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1000 BP; 340 A; 170 C; 231 G; 259 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1000;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 843 TTCTGAGAA 851

CP 33 ttctcagaa 25
ID US-08-459-818-17 STANDARD; DNA: UNC; 1021 BP.
AC xxxxxx
DE Sequence 17, Application US/08459818
CC Sequence 17, Application US/08459818
CC Patent No. 5851795
CC GENERAL INFORMATION:
CC APPLICANT: Linsley, Peter S.
CC APPLICANT: Ledbetter, Jeffrey A.
CC APPLICANT: Dangle, Nitin K.
CC APPLICANT: Brady, William
CC TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
CC NUMBER OF SEQUENCES: 27
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Merchant & Gould
CC STREET: 11150 Santa Monica Blvd., Suite 400
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90025
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Fastseq 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/459,818
CC FILING DATE: 02-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Adriano, Sarah B.
CC REGISTRATION NUMBER: 34,470
CC REFERENCE/DOCKET NUMBER: 30436.35US02
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 310-445-1140
CC TELEFAX: 310-445-9031
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1021 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1021 BP; 260 A; 246 C; 270 G; 245 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1021;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 529 TTCTGAGAA 537
OY 25 ttctgagaa 33
RESULT 125
ID US-08-247-809A-1 STANDARD; DNA: UNC; 1040 BP.
AC xxxxxx
DE Sequence 1, Application US/08247809A
CC Sequence 1, Application US/08247809A
CC Patent No. 5568823
CC GENERAL INFORMATION:
CC APPLICANT: Peter H. Schreier; Klaus Stenzel; Gunter Adam;
CC APPLICANT: Edgar Maiss
CC TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: SPRUNG HORN KRAMER & WOODS

CC STREET: 660 White Plains Road
CC CITY: Tarrytown
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10591-5144
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
CC OPERATING SYSTEM: DOS
CC SOFTWARE: WordPerfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/247,809A
CC FILING DATE: May 23, 1994
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: P 43 178 45.6 (Germany)
CC FILING DATE: May 28, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: KURT G. BRISCOE
CC REGISTRATION NUMBER: 33,141
CC REFERENCE/DOCKET NUMBER: Bayer 9049-KGB
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (914) 332-1700
CC TELEFAX: (914) 332-1844
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1040 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 36..989
CC SEQUENCE 1040 BP: 369 A; 186 C; 199 G; 286 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 1; Length 1040;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 545 TTCTGAGAA 553
Oy 25 ttctgagaa 33
RESULT 126
ID US-08-843-521-5 STANDARD; DNA; UNC; 1047 BP.
AC xxxxxx
DE Sequence 5, Application US/08843521
CC Sequence 5, Application US/08843521
CC Patent No. 5753479
CC GENERAL INFORMATION:
CC APPLICANT: Lawlor, Elizabeth
CC TITLE OF INVENTION: No. 5753479e1 Compounds
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Smithkline Beecham Corporation
CC STREET: 709 Swedeland Road
CC CITY: King of Prussia
CC STATE: PA
CC COUNTRY: USA
CC ZIP: 19406-0939
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC OPERATING SYSTEM: IBM Compatible
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/843,521
CC FILING DATE: 18-APR-1997

CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 9607993.4
CC FILING DATE: 18-APR-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GIMML, Edward R.
CC REGISTRATION NUMBER: 38,891
CC REFERENCE/DOCKET NUMBER: P31457-6
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 610-270-4478
CC TELEFAX: 610-270-5090
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1047 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: Genomic DNA
SQ SEQUENCE 1047 BP: 284 A; 221 C; 261 G; 281 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 2; Length 1047;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1027 TTCTGAGAA 1035
Op 33 ttctgagaa 25
RESULT 127
ID US-07-940-605A-11 STANDARD; DNA; UNC; 1060 BP.
AC xxxxxx
DE Sequence 11, Application US/07940605A
CC Sequence 11, Application US/07940605A
CC Patent No. 5540926
CC GENERAL INFORMATION:
CC APPLICANT: ARUFFO, ALEANDRO
CC APPLICANT: HOLLENBAUGH, DIANE
CC APPLICANT: LEDBETTER, JEFFREY A.
CC TITLE OF INVENTION: SOLUBLE LIGANDS FOR CD40
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/940,605A
CC FILING DATE: 04-SEP-1992
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 5624-184
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1060 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1060 BP: 196 A; 549 C; 294 G; 221 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1060;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1049 TTCTCAGAA 1057
|||||
33 ttctcagaa 25

RESULT 128
ID PCT-US96-03940-4 STANDARD; DNA; UNC; 1090 BP.
AC xxxxxx

DE Sequence 4, Application PC/TUS9603940
CC Sequence 4, Application PC/TUS9603940
CC GENERAL INFORMATION:

CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC TITLE OF INVENTION: REST Protein and DNA

CC NUMBER OF SEQUENCES: 29

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Dechert Price & Rhoads

CC STREET: P.O. Box 5218

CC CITY: Princeton

CC STATE: New Jersey

CC COUNTRY: USA

CC ZIP: 08543-5218

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

CC COMPUTER: IBM-compatible
CC OPERATING SYSTEM: DOS 5.0

CC SOFTWARE: WordPerfect

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03940

CC FILING DATE: March 23, 1995

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Allen Bloom

CC REGISTRATION NUMBER: 29,135

CC REFERENCE/DOCKET NUMBER: 317743-101

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214

CC TELEFAX: (609) 520-3259

CC INFORMATION FOR SEQ ID NO: 4:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1090 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double

CC TOPOLOGY: linear

CC MOLECULE TYPE: CDNA to mRNA

CC HYPOTHETICAL: no

CC ANTI-SENSE: no

CC ORIGINAL SOURCE:

CC ORGANISM: Human

CC CELL LINE: HeLa

CC IMMEDIATE SOURCE:

CC LIBRARY: CDNA

CC PUBLICATION INFORMATION:

CC AUTHORS: Chong, Jayhong A., Tapla-Ramirez Jos, Toledo-

CC AUTHORS: Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,

CC AUTHORS: Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

CC *TITLE: REST: A Mammalian Silencer Protein that Restricts

CC TITLE: Sodium Channel Gene Expression to Neurons

CC JOURNAL: Cell

CC VOLUME: 80

CC DATE: March 24, 1995

CC RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 TO 1090

CC SEQUENCE 1090 BP: 433 A; 206 C; 219 G; 232 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1090;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 304 TTCTCAGAA 312
|||||
33 ttctcagaa 25

RESULT 129
ID US-08-592-126-95 STANDARD; DNA; UNC; 1093 BP.
AC xxxxxx

DE Sequence 95, Application US/08592126
CC Sequence 95, Application US/08592126
CC Patent No. 5821091

CC GENERAL INFORMATION:

CC APPLICANT: Gregory Dolganov

CC TITLE OF INVENTION: Transcripts Encoding Immunomodulatory

CC TITLE OF INVENTION: Polypeptides

CC NUMBER OF SEQUENCES: 151

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Dehlinger & Associates

CC STREET: 350 Cambridge Avenue, Suite 250

CC CITY: Palo Alto

CC STATE: CA

CC COUNTRY: USA

CC ZIP: 94306

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent in Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/592,126

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Sholtz, Charles K.

CC REGISTRATION NUMBER: 38,615

CC REFERENCE/DOCKET NUMBER: 4600-0111

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (415) 324-0880

CC TELEFAX: (415) 324-0960

CC INFORMATION FOR SEQ ID NO: 95:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1093 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double

CC TOPOLOGY: linear

CC MOLECULE TYPE: CDNA to mRNA

CC HYPOTHETICAL: NO

CC ANTI-SENSE: NO

CC ORIGINAL SOURCE:

CC INDIVIDUAL ISOLATE: G164con.seq

CC SEQUENCE 1093 BP: 306 A; 234 C; 249 G; 304 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1093;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 775 TTCTCAGAA 783
|||||
25 ttctcagaa 33

RESULT 130
ID US-08-196-350-2 STANDARD; DNA; UNC; 1100 BP.
AC xxxxxx

DE Sequence 2, Application US/08196350
CC Sequence 2, Application US/08196350
CC Patent No. 5585099

CC GENERAL INFORMATION:
CC APPLICANT: Richards, Sue
CC APPLICANT: Kaplan, Joanne
CC APPLICANT: Mosicki, Richard
CC TITLE OF INVENTION: PROLACTIN AS ADJUVANT
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Brad Salcedo
CC STREET: One Kendall Square
CC CITY: Cambridge
CC STATE: MA
CC COUNTRY: U.S.A.
CC ZIP: 02139
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/196,350
CC FILING DATE:
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Gosz, William G
CC REGISTRATION NUMBER: 27,787
CC REFERENCE/DOCKET NUMBER: GEN 4-1.0
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/2527868
CC TELEFAX: 617/3747225
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1100 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC SEQUENCE 1100 BP: 324 A; 282 C; 202 G; 292 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 1; Length 1100;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 977 TTCTGAGAA 985
OY 25 ttctgagaa 33
RESULT 131
ID US-08-041-648-4 STANDARD; DNA; UNC; 1106 BP.
AC xxxxxx
DE Sequence 4, Application US/08041648
CC Sequence 4, Application US/08041648
CC Patent No. 5486463
CC GENERAL INFORMATION:
CC APPLICANT: Lesslauer, Werner
CC APPLICANT: L. tscher, Hansruedi
CC APPLICANT: St ber, Dietrich
CC TITLE OF INVENTION: TNF-MUTAINS
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: New Jersey
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/041,648
CC FILING DATE: 1-APR-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP 92810249.0
CC FILING DATE: 2-APR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Roseman, Catherine R.
CC REGISTRATION NUMBER: 34240
CC REFERENCE/DOCKET NUMBER: RAN 4105/147
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 235-6208
CC TELEFAX: (201) 235-3500
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1106 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: N-terminal
CC ORIGINAL SOURCE:
CC ORGANISM: Plasmid pDS56/RBSII, SphI-THFalpha(Trp32)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 994..1104
SQ SEQUENCE 1106 BP: 305 A; 256 C; 247 G; 298 T; 0 OTHER.
DB 379 TTCTGAGAA 387
OY 25 ttctgagaa 33
Query Match 100.0%; Score 9; DB 1; Length 1106;
Best Local Similarity 100.0%; Pred. No. 3,96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DE Sequence 6, Application US/08041648
CC Sequence 6, Application US/08041648
CC Patent No. 5486463
CC GENERAL INFORMATION:
CC APPLICANT: Lesslauer, Werner
CC APPLICANT: L. tscher, Hansruedi
CC APPLICANT: St ber, Dietrich
CC TITLE OF INVENTION: TNF-MUTAINS
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: New Jersey
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/041,648
CC FILING DATE: 1-APR-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP 92810249.0
CC FILING DATE: 2-APR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Roseman, Catherine R.

CC REGISTRATION NUMBER: 34240
CC REFERENCE/DOCKET NUMBER: RAN 4105/147
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 235-6208
CC TELEFAX: (201) 235-3500
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1106 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: N-terminal
CC ORIGINAL SOURCE:
CC ORGANISM: Plasmid pDS56/RBSII, Sphi-TNFalpha(Ser29)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 994..1104
CC SEQUENCE 1106 BP; 305 A; 258 C; 246 G; 297 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 1106;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 379 TTCTGAGAA 387
QY 25 ttctgagaa 33
RESULT 133
ID US-08-041-648-8 STANDARD; DNA; UNC; 1106 BP.
AC xxxxxx
DE Sequence 8, Application US/08041648
CC Sequence 8, Application US/08041648
CC Patent No. 5486463
CC GENERAL INFORMATION:
CC APPLICANT: Lesslauer, Werner
CC APPLICANT: L. tescher, Hansruedi
CC APPLICANT: St Ber, Dietrich
CC TITLE OF INVENTION: TNF-MUTAINS
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
CC STREET: 340 Kingsland Street
CC CITY: Nutley
CC STATE: New Jersey
CC COUNTRY: U.S.A.
CC ZIP: 07110-1199
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/041.648
CC FILING DATE: 1-APR-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP 92810249.0
CC FILING DATE: 2-APR-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Roseman, Catherine R.
CC REGISTRATION NUMBER: 34240
CC REFERENCE/DOCKET NUMBER: RAN 4105/147
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 235-6208
CC TELEFAX: (201) 235-3500
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1106 base pairs

CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: N-terminal
CC ORIGINAL SOURCE:
CC ORGANISM: Plasmid pDS56/RBSII, Sphi-TNFalpha(Ser29Trp32)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 994..1104
CC SEQUENCE 1106 BP; 305 A; 257 C; 246 G; 298 T; 0 OTHER.
SO
Query Match 100.0%; Score 9; DB 1; Length 1106;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 379 TTCTGAGAA 387
QY 25 ttctgagaa 33
RESULT 134
ID US-08-630-918-57 STANDARD; DNA; UNC; 1110 BP.
AC xxxxxx
DE Sequence 57, Application US/08630918
CC Sequence 57, Application US/08630918
CC Patent No. 5631350
CC GENERAL INFORMATION:
CC APPLICANT: Yoshinobu OKUNO et al.
CC TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY
CC NUMBER OF SEQUENCES: 58
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Wenderoth, Lind & Ponack
CC STREET: 805 Fifteenth Street, N.W., #700
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: U.S.A.
CC ZIP: 20005
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: Wordperfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/630.918
CC FILING DATE: April 5, 1996
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/229,781
CC FILING DATE: April 19, 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/054,016
CC FILING DATE: April 29, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Matteo M. Cheek, Jr.
CC REGISTRATION NUMBER: 33,367
CC REFERENCE/DOCKET NUMBER:
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-371-8850
CC TELEFAX:
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 57:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1110 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to genomic RNA
CC ORIGINAL SOURCE:
CC ORGANISM: A2/Alch1/2/68

SO SEQUENCE 1110 BP; 374 A; 216 C; 258 G; 262 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1110;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 619 TTCTCAGAA 627
CP 33 ttctcagaa 25

RESULT 135
ID US-08-229-781-57 STANDARD; DNA; UNC; 1110 BP.
AC xxxxxx
DT Sequence 57, Application US/08229781
DE Sequence 57, Application US/08229781
CC Patent No. 5589174
CC GENERAL INFORMATION:
CC APPLICANT: Yoshinobu OKUNO et al.
CC TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY
CC NUMBER OF SEQUENCES: 58
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Menderoth, Lind & Ponack
CC STREET: 805 Fifteenth Street, N.W., #700
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: U.S.A.
CC ZIP: 20005
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: Wordperfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/229,781
CC FILING DATE: April 19, 1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/054,016
CC FILING DATE: April 29, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Warren M. Cheek, Jr.
CC REGISTRATION NUMBER: 33,367
CC REFERENCE/DOCKET NUMBER:
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-371-8850
CC TELEFAX:
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 57:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1110 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to genomic RNA
CC HYPOTHETICAL:
CC ANTI-SENSE:
CC FRAGMENT TYPE:
CC ORIGINAL SOURCE:
CC ORGANISM: A2/Aichi/2/68
CC STRAIN:
CC INDIVIDUAL ISOLATE:
CC DEVELOPMENTAL STAGE:
CC HAPLOTYPE:
CC TISSUE TYPE:
CC CELL TYPE:
CC CELL LINE:
CC ORGANELLE:
CC IMMEDIATE SOURCE:
CC LIBRARY:
CC CLONE:
CC POSITION IN GENOME:

CC CHROMOSOME/SEGMENT:
CC MAP POSITION:
CC UNITS:
CC FEATURE:
CC NAME/KEY:
CC LOCATION:
CC IDENTIFICATION METHOD:
CC OTHER INFORMATION:
CC PUBLICATION INFORMATION:
CC AUTHORS:
CC TITLE:
CC JOURNAL:
CC VOLUME:
CC ISSUE:
CC PAGES:
CC DATE:
CC DOCUMENT NUMBER:
CC FILING DATE:
CC PUBLICATION DATE:
CC RELEVANT RESIDUES IN SEQ ID NO:
SO SEQUENCE 1110 BP; 374 A; 216 C; 258 G; 262 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1110;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 619 TTCTCAGAA 627
CP 33 ttctcagaa 25

RESULT 136
ID PCT-US94-13187-1 STANDARD; DNA; UNC; 1129 BP.
AC xxxxxx
DT Sequence 1, Application PC/TUS9413187
DE Sequence 1, Application PC/TUS9413187
CC GENERAL INFORMATION:
CC APPLICANT: WEI, ET AL.
CC TITLE OF INVENTION: Human MUT2
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,
CC ADDRESSEE: CECCHI, STEWART & OLSTEIN
CC STREET: 6 BECKER FARM ROAD
CC CITY: ROSELAND
CC STATE: NEW JERSEY
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 INCH DISKETTE
CC COMPUTER: IBM PS/2
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WORD PERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/13187
CC FILING DATE: Submitted herewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FERRARO, GREGORY D.
CC REGISTRATION NUMBER: 36,134
CC REFERENCE/DOCKET NUMBER: 325800-245
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-994-1700
CC TELEFAX: 201-994-1744
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1129 BASE PAIRS
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: SINGLE

CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1129 BP; 333 A; 244 C; 258 G; 294 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1129;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 TTCTCAGAA 154
|||
Cp 33 ttctcagaa 25

RESULT 137
ID US-08-470-261-1 STANDARD; DNA; UNC; 1129 BP.
AC xxxxxx

DE Sequence 1, Application US/08470261
CC Sequence 1, Application US/08470261
CC Patent No. 5695980
CC GENERAL INFORMATION:
CC APPLICANT: WEI, ET AL.
CC TITLE OF INVENTION: Human MULT2
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CC ADDRESSEE: CECCHI, STEWART & OLSTEIN
CC STREET: 6 BECKER FARM ROAD
CC CITY: ROSELAND
CC STATE: NEW JERSEY
CC COUNTRY: USA
CC ZIP: 07068

COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 INCH DISKETTE
CC COMPUTER: IBM PS/2
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WORD PERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/470,261
CC FILING DATE: June 6 1995

CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/13187
CC FILING DATE: 15 NOV 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FERRARO, GREGORY D.
CC REGISTRATION NUMBER: 36,134
CC REFERENCE/DOCKET NUMBER: 325800-372
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-994-1700
CC TELEFAX: 201-994-1744
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1129 BASE PAIRS
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: SINGLE
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1129 BP; 333 A; 244 C; 258 G; 294 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 2; Length 1129;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 TTCTCAGAA 154
|||
Cp 33 ttctcagaa 25

RESULT 138
ID US-08-487-810-1 STANDARD; DNA; UNC; 1146 BP.
AC xxxxxx

DE Sequence 1, Application US/08487810
CC Sequence 1, Application US/08487810
CC Patent No. 5618695
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: DNA ENCODING HEM-1, A GENE EXPRESSED BY
CC TITLE OF INVENTION: SCLEROSING HEMANGIOMA CELLS
CC NUMBER OF SEQUENCES: 2
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/487,810
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1146 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1146 BP; 459 A; 193 C; 238 G; 256 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1146;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 178 TTCTCAGAA 186
|||
Cp 33 ttctcagaa 25

RESULT 139
ID US-08-456-104-3 STANDARD; DNA; UNC; 1151 BP.
AC xxxxxx

DE Sequence 3, Application US/08456104
CC Sequence 3, Application US/08456104
CC Patent No. 5861310
CC GENERAL INFORMATION:
CC APPLICANT: Freeman, Gordon J.
CC APPLICANT: Nadler, Lee M.
CC APPLICANT: Gray, Gary S.
CC TITLE OF INVENTION: TUMOR CELLS MODIFIED TO EXPRESS B7-2 AND B7-3 WITH INCRE
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: LAHYE & COCKFIELD
CC STREET: 60 State Street, Suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109

COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/456,104
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/101,624;
CC FILING DATE: 26-JUL-1993;
CC APPLICATION NUMBER: 08/109,393;
CC APPLICATION NUMBER: 19-AUG-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mandregouras, Amy E.
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: RPI-008
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400

CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1151 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 99..1028
CC SEQUENCE 1151 BP: 353 A; 258 C; 255 G; 285 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 3; Length 1151;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1091 TTCTCAGAA 1099
Cp 33 ttctcagaa 25
RESULT 140
ID PCT-US95-02576-20 STANDARD; DNA; UNC; 1151 BP.
AC xxxxxx
DE Sequence 20, Application PC/TUS9502576
CC Sequence 20, Application PC/TUS9502576
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
CC NUMBER OF SEQUENCES: 65
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02576
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/205,697
CC FILING DATE: 02-Mar-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mandagouras, Amy E.
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: BWI-120CPCPC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1151 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 99..1025
CC SEQUENCE 1151 BP: 353 A; 258 C; 255 G; 285 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 4; Length 1151;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1091 TTCTCAGAA 1099
Cp 33 ttctcagaa 25
RESULT 141
ID PCT-US95-03032-2 STANDARD; DNA; UNC; 1176 BP.
AC xxxxxx
DE Sequence 2, Application PC/TUS9503032
CC Sequence 2, Application PC/TUS9503032
CC GENERAL INFORMATION:
CC APPLICANT: Repligen Corporation
CC APPLICANT: the Trustees of Boston University
CC TITLE OF INVENTION: ANTIBODIES TO INTERLEUKIN-8 RECEPTORS AND
CC METHODS OF USE
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5" Diskette, 1.44 MB
CC COMPUTER: IBM PS/2 Model 502 or 55SX
CC OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
CC SOFTWARE: WordPerfect (Version 5.0)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/03032
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/237,937
CC FILING DATE: 02-MAY-94
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/210,250
CC FILING DATE: 15-MAR-94
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/803,842
CC FILING DATE: 09-DEC-91
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/726,606
CC FILING DATE: 09-JUL-91
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/685,101
CC FILING DATE: 10-APR-91
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fasse, J. Peter
CC REGISTRATION NUMBER: 32,983
CC REFERENCE/DOCKET NUMBER: 04766/015W01
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1176
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 1176 BP: 247 A; 333 C; 279 G; 317 T; 0 OTHER.
SQ
Query Match 100.0%; Score 9; DB 4; Length 1176;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1108 TTCTCAGAA 1116
Cp 33 ttctcagaa 25

RESULT 142
ID US-08-493-638-5 STANDARD; DNA; UNC; 1185 BP.
AC xxxxxx
DE Sequence 5, Application US/08493638
CC Sequence 5, Application US/08493638
CC Patent No. 5824299
CC GENERAL INFORMATION:
CC APPLICANT: Luster, Andrew
CC APPLICANT: Leder, Philip
CC TITLE OF INVENTION: MODULATION OF ENDOTHELIAL CELL
CC TITLE OF INVENTION: PROLIFERATION
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/493,638
CC FILING DATE: 22-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/542-5070
CC TELEFAX: 617/542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1185 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 80..374
CC SEQUENCE 1185 BP; 389 A; 234 C; 210 G; 352 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1185;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 731 TTCACGAA 739
Cp 33 ttccagaa 25
RESULT 143
ID PCT-US95-07743-6 STANDARD; DNA; UNC; 1188 BP.
AC xxxxxx
DE Sequence 6, Application PC/TUS9507743
CC Sequence 6, Application PC/TUS9507743
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC
CC TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM AND PROMOTERS AND TERMINATORS FC
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Novo Nordisk of North America, Inc.
CC STREET: 405 Lexington Avenue, 64th Floor

CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10174-6401
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/07743
CC FILING DATE: 15-June-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/269,449
CC FILING DATE: 30-June-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/404,678
CC FILING DATE: 15-March-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Agtis Dr., Cheryl H.
CC REGISTRATION NUMBER: 34,086
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-867-0123
CC TELEFAX: 212-878-9655
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1188 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 1188 BP; 366 A; 229 C; 252 G; 341 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1031 TTCACGAA 1039
Cp 33 ttccagaa 25
RESULT 144
ID US-08-921-426-6 STANDARD; DNA; UNC; 1188 BP.
AC xxxxxx
DE Sequence 6, Application US/08921426
CC Sequence 6, Application US/08921426
CC Patent No. 5837847
CC GENERAL INFORMATION:
CC APPLICANT: Royer, John C
CC APPLICANT: Moyer, Donna L
CC APPLICANT: Yoder, Wendy T
CC APPLICANT: Shuster, Jeffrey R
CC TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC
CC TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: No. 5837847 of No. 5837847th America, Inc.
CC STREET: 405 Lexington Avenue, 64th Floor
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10174-6401
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/921,426
CC FILING DATE: 29-AUG-1997

CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/456,433
CC FILING DATE: 01-JUN-1995
CC APPLICATION NUMBER: US 08/404,678
CC FILING DATE: 15-MAR-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Agilis Dr., Cheryl H.
CC REGISTRATION NUMBER: 34,086
CC REFERENCE/DOCKET NUMBER: 4216, 010-US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-867-0123
CC TELEFAX: 212-878-9655
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1188 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SO SEQUENCE 1188 BP; 366 A; 229 C; 252 G; 341 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1031 TTCTCAGAA 1039
|||||
Cp 33 ttctcagaa 25

RESULT 145
ID US-08-592-126-56 STANDARD; DNA: UNC; 1189 BP.
AC xxxxxx

DE Sequence 56, Application US/08592126
CC Sequence 56, Application US/08592126
CC Patent No. 5821091
CC GENERAL INFORMATION:
CC APPLICANT: Gregory Dolganov
CC TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
CC NUMBER OF SEQUENCES: 151
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Dehlinger & Associates
CC STREET: 350 Cambridge Avenue, Suite 250
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/592,126
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sholtz, Charles K.
CC REGISTRATION NUMBER: 38,615
CC REFERENCE/DOCKET NUMBER: 4600-0111
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 324-0880
CC TELEFAX: (415) 324-0960
CC INFORMATION FOR SEQ ID NO: 56:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1189 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO

CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC INDIVIDUAL ISOLATE: Tc1.seq
SO SEQUENCE 1189 BP; 307 A; 281 C; 235 G; 366 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1173 TTCTCAGAA 1181
|||||
Cp 33 ttctcagaa 25

RESULT 146
ID US-08-864-799-2 STANDARD; DNA: UNC; 1196 BP.
AC xxxxxx

DE Sequence 2, Application US/08864799
CC Sequence 2, Application US/08864799
CC Patent No. 5858714
CC GENERAL INFORMATION:
CC APPLICANT: Hillman, Jennifer L.
CC APPLICANT: Corley, Neil C.
CC APPLICANT: Shah, Parvi
CC TITLE OF INVENTION: HUMAN METAXIN PROTEIN
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/864,799
CC FILING DATE: Filed Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0312 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-845-4166
CC TELEFAX: 415-845-0555
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1196 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: ENDCNOT01
CC CLONE: 2134944
SO SEQUENCE 1196 BP; 338 A; 221 C; 275 G; 362 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1196;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 830 TTCTCAGAA 838
|||||
Cp 25 ttctcagaa 33

RESULT 147
ID US-08-434-881-1 STANDARD; DNA; UNC; 1221 BP.
AC xxxxxx
DT
Sequence 1, Application US/08434881
CC Sequence 1, Application US/08434881
CC Patent No. 5804376
CC GENERAL INFORMATION:
CC APPLICANT: Braxton, Scott M.
CC APPLICANT: Wilep, Craig G.
CC APPLICANT: Diep, Dinh
CC TITLE OF INVENTION: Pancreas-Derived Serpin
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Inocyte Pharmaceuticals, Inc.
CC STREET: 3330 Hillview Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/434,881
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Luther, Barbara J.
CC REGISTRATION NUMBER: 33954
CC REFERENCE/DOCKET NUMBER: PF0035 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-852-0195
CC TELEFAX: 415-852-0195
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1221 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC IMMEDIATE SOURCE:
CC LIBRARY: Pancreas
CC CLONE: 222689
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..1221
SQ SEQUENCE 1221 BP; 412 A; 210 C; 249 G; 337 T; 13 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1221;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 659 TTCTGAGAA 667
OY 25 ttctgagaa 33

CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Grace L. Bonner, Monsanto CO., BBAF
CC STREET: 700 Chesterfield Parkway No. 538653th
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: United States of America
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/334,639
CC FILING DATE:
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bonner, Grace L.
CC REGISTRATION NUMBER: 32,963
CC REFERENCE/DOCKET NUMBER: 38-21(10560)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314)537-7286
CC TELEFAX: (314)537-6047
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1226 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1226 BP; 378 A; 236 C; 203 G; 409 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1226;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1135 TTCTCAGAA 1143
Cp 33 ttctcagaa 25
RESULT 149
ID US-08-090-523-26 STANDARD; DNA; UNC; 1226 BP.
AC xxxxxx
DT
Sequence 26, Application US/08090523
CC Sequence 26, Application US/08090523
CC Patent No. 5498830
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Stark, David M.
CC TITLE OF INVENTION: Enhanced Starch Biosynthesis
CC NUMBER OF SEQUENCES: 51
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF
CC STREET: 700 Chesterfield Parkway No. 5498830th
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/090,523
CC FILING DATE: 19930712
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/709663

CC FILING DATE: 07-JUN-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/539763
CC FILING DATE: 18-JUN-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bonner, Grace L.
CC REGISTRATION NUMBER: 32,963
CC REFERENCE/DOCKET NUMBER: 38-21(10559)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314) 537-7286
CC TELEFAX: (314) 537-6047
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1226 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1226 BP: 378 A; 236 C; 203 G; 409 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1226;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1135 TTCTCAGAA 1143
Cp 33 ttctcagaa 25

RESULT 150
ID US-08-398-627-26 STANDARD; DNA; UNC; 1226 BP.
AC xxxxxx
DE Sequence 26, Application US/08398627
DE Sequence 26, Application US/08398627
DE Patent No. 5608149
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Stark, David M.
CC TITLE OF INVENTION: Enhanced starch biosynthesis
CC NUMBER OF SEQUENCES: 51
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Grace L. Bonner, Monsanto Co. BBAF
CC STREET: 700 Chesterfield Parkway No. 5608149th
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/398,627
CC FILING DATE: 03-MAR-1995
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/090,523
CC FILING DATE: 12-JUL-1993
CC APPLICATION NUMBER: US 07/709663
CC FILING DATE: 07-JUN-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/539763
CC FILING DATE: 18-JUN-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bonner, Grace L.
CC REGISTRATION NUMBER: 32,963
CC REFERENCE/DOCKET NUMBER: 38-21(10559)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314) 537-7286
CC TELEFAX: (314) 537-6047

CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1226 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 1226 BP: 378 A; 236 C; 203 G; 409 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1226;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1135 TTCTCAGAA 1143
Cp 33 ttctcagaa 25

RESULT 151
ID US-09-031-485-18 STANDARD; DNA; UNC; 1227 BP.
AC xxxxxx
DE Sequence 18, Application US/09031485
DE Sequence 18, Application US/09031485
DE Patent No. 5824306
CC GENERAL INFORMATION:
CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESSES:
CC ADDRESS: Carol Talkington Verser, Ph.D.
CC ADDRESS: Heska Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/031,485
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429
CC FILING DATE: 24-APR-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/493-7272
CC TELEFAX: 970/484-9505
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1227 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1227 BP: 320 A; 288 C; 265 G; 353 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1227;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 916 TTCTGAGAA 924

QY 25 ttcctgagaa 33

RESULT 152
ID US-08-847-429A-18 STANDARD; DNA; UNC; 1227 BP.
AC xxxxxx

DE Sequence 18, Application US/08847429A
CC Sequence 18, Application US/08847429A
CC Patent No. 5827692
CC GENERAL INFORMATION:

CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hesk Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429A
CC FILING DATE: 24-Apr-1997

CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/493-7272
CC TELEFAX: 970/484-9505

CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1227 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1227 BP; 320 A; 288 C; 265 G; 353 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1227;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 916 TTCTGAGAA 924
OY 25 ttcctgagaa 33

RESULT 153
ID US-09-031-485-17 STANDARD; DNA; UNC; 1227 BP.
AC xxxxxx

DE Sequence 17, Application US/09031485
CC Sequence 17, Application US/09031485
CC Patent No. 5824306
CC GENERAL INFORMATION:

CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC TITLE OF INVENTION: USES THEREOF

CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Hesk Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/031,485
CC FILING DATE:

CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429
CC FILING DATE: 24-Apr-1997
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/493-7272
CC TELEFAX: 970/484-9505

CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1227 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1227 BP; 353 A; 265 C; 288 G; 320 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1227;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 304 TTCTGAGAA 312
Cp 33 ttcctgagaa 25

RESULT 154
ID US-08-847-429A-17 STANDARD; DNA; UNC; 1227 BP.
AC xxxxxx

DE Sequence 17, Application US/08847429A
CC Sequence 17, Application US/08847429A
CC Patent No. 5827692
CC GENERAL INFORMATION:

CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Hesk Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429A
CC FILING DATE: 24-APR-1997
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/484-9505
CC TELEFAX: 970/493-7272
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1227 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1227 BP: 353 A; 265 C; 288 G; 320 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1227;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 304 TTCTGAGAA 312
|||||
Cp 33 ttctcgaa 25

RESULT 155
ID US-09-031-485-16 STANDARD; DNA; UNC; 1228 BP.
AC xxxxxx
DE Sequence 16, Application US/09031485
CC Sequence 16, Application US/09031485
CC Patent No. 5824306
CC GENERAL INFORMATION:
CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC TITLE OF INVENTION: USES THEREOF
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Heeka Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/031,485
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429
CC FILING DATE: 24-APR-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/484-9505
CC TELEFAX: 970/493-7272
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1228 nucleotides

CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1228 BP: 320 A; 288 C; 266 G; 353 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 917 TTCTGAGAA 925
|||||
QY 25 ttctcgaa 33

RESULT 156
ID US-08-847-429A-16 STANDARD; DNA; UNC; 1228 BP.
AC xxxxxx
DE Sequence 16, Application US/08847429A
CC Sequence 16, Application US/08847429A
CC Patent No. 5827692
CC GENERAL INFORMATION:
CC APPLICANT: Tang, Liang
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC TITLE OF INVENTION: USES THEREOF
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Heeka Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429A
CC FILING DATE: 24-APR-1997
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/484-9505
CC TELEFAX: 970/493-7272
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1228 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1228 BP: 320 A; 288 C; 266 G; 353 T; 1 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 917 TTCTGAGAA 925
|||||
QY 25 ttctcgaa 33

RESULT 157
ID US-09-031-485-14 STANDARD; DNA; UNC; 1228 BP.

AC xxxxxx
DE Sequence 14, Application US/09031485
CC Sequence 14, Application US/09031485
CC Patent No. 5824306
CC GENERAL INFORMATION:
CC APPLICANT: Tang, E. Scot
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC TITLE OF INVENTION: USES THEREOF
CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Heska Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/031,485
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429
CC FILING DATE: 24-APR-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/493-7272
CC TELEFAX: 970/484-9505
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1228 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..1227
SQ SEQUENCE 1228 BP; 353 A; 266 C; 288 G; 320 T; 1 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 304 TTCTCAGAA 312
|||||
Cp 33 ttctcagaa 25
RESULT 158
ID US-08-847-429A-14 STANDARD; DNA; UNC; 1228 BP.
AC xxxxxx
DE Sequence 14, Application US/08847429A
CC Sequence 14, Application US/08847429A
CC Patent No. 5827692
CC GENERAL INFORMATION:
CC APPLICANT: Tang, E. Scot
CC APPLICANT: Blehm, E. Scot
CC TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
CC TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
CC TITLE OF INVENTION: USES THEREOF

CC NUMBER OF SEQUENCES: 85
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carol Talkington Verser, Ph.D.
CC ADDRESSEE: Heska Corporation
CC STREET: 1825 Sharp Point Drive
CC CITY: Fort Collins
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80525
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: Windows 95
CC SOFTWARE: Wordperfect for Windows, Version 7.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/847,429A
CC FILING DATE: 24-APR-1997
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Verser, Carol Talkington
CC REGISTRATION NUMBER: 37,459
CC REFERENCE/DOCKET NUMBER: HW-5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 970/493-7272
CC TELEFAX: 970/484-9505
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1228 nucleotides
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..1227
SQ SEQUENCE 1228 BP; 353 A; 266 C; 288 G; 320 T; 1 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1228;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 304 TTCTCAGAA 312
|||||
Cp 33 ttctcagaa 25
RESULT 159
ID PCT-US93-03077-4 STANDARD; DNA; UNC; 1242 BP.
AC xxxxxx
DE Sequence 4, Application PC/TUS9303077
CC Sequence 4, Application PC/TUS9303077
CC GENERAL INFORMATION:
CC APPLICANT: Board of Regents, The University of Texas System
CC APPLICANT: Gaynor, Richard B.
CC APPLICANT: Wu, Foon Kin
CC TITLE OF INVENTION: PROTEIN CELLULAR FACTOR USEFUL FOR
CC TITLE OF INVENTION: REGULATING GENE EXPRESSION
CC NUMBER OF SEQUENCES: 7
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/03077

CC FILING DATE: 19930331
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/862.025
CC FILING DATE: April 2, 1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kammerer, Patricia A.
CC REGISTRATION NUMBER: 29,775
CC REFERENCE/DOCKET NUMBER: UTPD270PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 713-787-1540
CC TELEFAX: 713-749-2679
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1242 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
SO SEQUENCE 1242 BP; 479 A; 215 C; 295 G; 253 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1242;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 TTCTCAGAA 23
|||||
0y 25 ttctcgaa 33

RESULT 160
ID PCT-US95-02576-12 STANDARD; DNA; UNC; 1261 BP.
AC xxxxxx
DE Sequence 12, Application PC/TUS9502576
CC Sequence 12, Application PC/TUS9502576
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules
CC NUMBER OF SEQUENCES: 65
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/02576
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/205,697
CC FILING DATE: 02-Mar-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mandragoras, Amy E.
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: BWT-120CPC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1261 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA

CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 194..1135
CC SEQUENCE 1261 BP; 358 A; 277 C; 279 G; 345 T; 2 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1261;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1201 TTCTCAGAA 1209
|||||
Cp 33 ttctcgaa 25

RESULT 161
ID US-07-794-731A-1 STANDARD; DNA; UNC; 1273 BP.
AC xxxxxx
DE Sequence 1, Application US/07794731A
CC Sequence 1, Application US/07794731A
CC Patent No. 534867
CC GENERAL INFORMATION:
CC APPLICANT: George Georgiou
CC APPLICANT: Joseph A. Francisco
CC APPLICANT: Charles F. Earnhart
CC TITLE OF INVENTION: Expression of Proteins on Bacterial
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: USA
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy Disk
CC COMPUTER: IBM PC Compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Wordperfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/794,731A
CC FILING DATE: 19911115
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kitchell, Barbara S.
CC REGISTRATION NUMBER: 33,928
CC REFERENCE/DOCKET NUMBER: GEOG:001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 320-7200
CC TELEFAX: (512) 474-7577
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1273 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
SO SEQUENCE 1273 BP; 339 A; 308 C; 332 G; 294 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1273;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 663 TTCTCAGAA 671
|||||
Cp 33 ttctcgaa 25

RESULT 162
ID US-07-794-731A-2 STANDARD; DNA; UNC; 1273 BP.
AC xxxxxx
DE Sequence 2, Application US/07794731A

CC FILING DATE: 06-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/844,055
CC FILING DATE: 18-APR-1997
CC APPLICATION NUMBER: 9607993.4
CC FILING DATE: 18-APR-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GIMMLI, Edward R
CC REGISTRATION NUMBER: 38,891
CC REFERENCE/DOCKET NUMBER: P31457-1/1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 610-270-4478
CC TELEFAX: 610-270-5090
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1290 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: Genomic RNA
SQ SEQUENCE 1290 BP: 367 A; 290 C; 305 G; 328 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 3; Length 1290;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1249 TTCGAGAA 1257
|||||
Cp 33 ttctcagaa 25
RESULT 165
ID PCT-US94-12912-3 STANDARD: DNA; UNC: 1296 BP.
AC xxxxxx
DT Sequence 3, Application PC/TUS9412912
CC Sequence 3, Application PC/TUS9412912
CC GENERAL INFORMATION:
CC APPLICANT: UNIVERSITY OF SOUTHERN CALIFORNIA
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: c/o Robbins, Berliner & Carson
CC STREET: 201 North Figueroa Street, Fifth Floor
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: U.S.A.
CC ZIP: 90012
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/12912
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Spitals, John P.
CC REGISTRATION NUMBER: 29,215
CC REFERENCE/DOCKET NUMBER: 1920-341
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (213) 977-1003
CC TELEFAX: (213) 977-1003
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1296 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: cDNA
SQ SEQUENCE 1296 BP: 372 A; 292 C; 303 G; 329 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 1296;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 442 TTCGAGAA 450
|||||
Qy 25 ttctcagaa 33
RESULT 166
ID US-07-679-052A-14 STANDARD: DNA; UNC: 1328 BP.
AC xxxxxx
DT Sequence 14, Application US/07679052A
DE Sequence 14, Application US/07679052A
CC Patent No. 5298400
CC GENERAL INFORMATION:
CC APPLICANT: WHITEFIELD, Peter L.
CC APPLICANT: RICHARDSON, Michael A.
CC APPLICANT: BUNN, Clive I.
CC TITLE OF INVENTION: RECOMBINANT PRODUCT
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/679,052A
CC FILING DATE: 19910506
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 16786/147 CHAC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1328 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: Other nucleic acid;
CC DESCRIPTION: Synthetic DNA oligonucleotide
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 15..1319
CC FEATURE:
CC NAME/KEY: sig_peptide
CC LOCATION: 15..71
CC OTHER INFORMATION: /product- "Synthetic signal
CC OTHER INFORMATION: peptide
SQ SEQUENCE 1328 BP: 377 A; 300 C; 317 G; 334 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1328;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1039 TTCGAGAA 1047
|||||

OY 25 ttctgagaa 33

RESULT 167
ID US-07-651-710A-1 STANDARD; DNA; UNC; 1338 BP.
AC xxxxxx
DE Sequence 1, Application US/07651710A
CC Sequence 1, Application US/07651710A
CC Patent No. 5362864
CC GENERAL INFORMATION:
CC APPLICANT: Chua, Nam-Hai
CC TITLE OF INVENTION: Trans-Activating Factor-1
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/651,710A
CC FILING DATE: 19910206
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 30,742
CC REFERENCE/DOCKET NUMBER: 3288-014
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 8698864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1338 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 9..803
CC SEQUENCE 1338 BP; 421 A; 244 C; 333 G; 340 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1338;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 395 TTCTGAGAA 403
OY 25 ttctgagaa 33

RESULT 168
ID US-07-679-052A-16 STANDARD; DNA; UNC; 1340 BP.
AC xxxxxx
DE Sequence 16, Application US/07679052A
CC Sequence 16, Application US/07679052A
CC Patent No. 5298400
CC GENERAL INFORMATION:
CC APPLICANT: WHITFIELD, Peter L.
CC APPLICANT: RICHARDSON, Michael A.
CC APPLICANT: BUNN, Clive L.
CC TITLE OF INVENTION: RECOMBINANT PRODUCT
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:

CC ADDRESS: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/679,052A
CC FILING DATE: 19910506
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 16786/147 CHAC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1340 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: Other nucleic acid;
CC DESCRIPTION: Synthetic DNA oligonucleotide
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 15..1331
CC FEATURE:
CC NAME/KEY: s1g-peptide
CC LOCATION: 15..86
CC OTHER INFORMATION: /product="a-1-antitrypsin signal"
CC SEQUENCE 1340 BP; 374 A; 308 C; 317 G; 341 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1340;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1051 TTCTGAGAA 1059
OY 25 ttctgagaa 33

RESULT 169
ID US-08-385-745-26 STANDARD; DNA; UNC; 1341 BP.
AC xxxxxx
DE Sequence 26, Application US/08385745
CC Sequence 26, Application US/08385745
CC Patent No. 5612209
CC GENERAL INFORMATION:
CC APPLICANT: Kling, Te Piao
CC TITLE OF INVENTION: Cloning and Recombinant Production of
CC TITLE OF INVENTION: Vespid Venom Phospholipases, and Immunological Therapies
CC NUMBER OF SEQUENCES: 27
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/385,745
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/08/031,400
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 3288-020
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1341 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 153..1052
SQ SEQUENCE 1341 BP; 473 A; 206 C; 239 G; 423 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 819 TTCTCAGAA 827
|||||
Cp 33 ttctcagaa 25

RESULT 170
ID PCT-US94-02629-26 STANDARD; DNA; UNC; 1341 BP.
AC xxxxxx

DE Sequence 26, Application PC/TUS9402629
CC Sequence 26, Application PC/TUS9402629
CC GENERAL INFORMATION:
CC APPLICANT: King, Te-piao
CC TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
CC TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC TITLE OF INVENTION: HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED THEREON
CC NUMBER OF SEQUENCES: 62
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Klauber & Jackson
CC STREET: 411 Hackensack Avenue
CC CITY: Hackensack
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07601
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/02629
CC FILING DATE: 10-MAR-1994
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/180,209
CC FILING DATE: 11-JAN-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/031,400
CC FILING DATE: 11-MAR-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jackson Esq., David A.

CC REGISTRATION NUMBER: 26,742
CC REFERENCE/DOCKET NUMBER: 600-1-074 PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201 487-5800
CC TELEFAX: 201 343-1684
CC TELEX: 133521
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1341 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 153..1052
SQ SEQUENCE 1341 BP; 473 A; 206 C; 239 G; 423 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 819 TTCTCAGAA 827
|||||
Cp 33 ttctcagaa 25

RESULT 171
ID US-08-180-209B-26 STANDARD; DNA; UNC; 1341 BP.
AC xxxxxx

DE Sequence 26, Application US/08180209B
CC Sequence 26, Application US/08180209B
CC Patent No. 3593877
CC GENERAL INFORMATION:
CC APPLICANT: King, Te-piao
CC TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
CC TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
CC TITLE OF INVENTION: HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
CC TITLE OF INVENTION: THEREON
CC NUMBER OF SEQUENCES: 62
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Klauber & Jackson
CC STREET: 411 Hackensack Avenue
CC CITY: Hackensack
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07601
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/180,209B
CC FILING DATE: 11-JAN-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/031,400
CC FILING DATE: 11-MAR-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jackson Esq., David A.
CC REGISTRATION NUMBER: 26,742
CC REFERENCE/DOCKET NUMBER: 600-1-074 CIP
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201 487-5800
CC TELEFAX: 201 343-1684
CC TELEX: 133521
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1341 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single

CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 153..1052
SQ SEQUENCE 1341 BP; 473 A; 206 C; 239 G; 423 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 819 TTCTCAGAA 827
|||||
Cp 33 ttctcagaa 25

RESULT 172
ID US-08-625-322-3 STANDARD; DNA; UNC; 1353 BP.
AC xxxxxx
DT
DE Sequence 3, Application US/08625322
CC Sequence 3, Application US/08625322
CC Patent No. 5804412
CC GENERAL INFORMATION:
CC APPLICANT: Gill, Gordon N.
CC APPLICANT: Korten, Richard C.
CC APPLICANT: Cadena, Deborah L.
CC TITLE OF INVENTION: Sorting Nexins and Methods of Using Same
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Campbell and Flores
CC STREET: 4370 La Jolla Village Drive, Suite 700
CC CITY: San Diego
CC STATE: California
CC COUNTRY: USA
CC ZIP: 92122
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/625,322
CC FILING DATE: 01-APR-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Campbell, Cathryn A.
CC REGISTRATION NUMBER: 31,815
CC REFERENCE/DOCKET NUMBER: P-UD 1955
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 535-9001
CC TELEFAX: (619) 535-8949
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1353 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..1353
SQ SEQUENCE 1353 BP; 461 A; 247 C; 304 G; 341 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1353;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 961 TTCTCAGAA 969
|||||
Cp 33 ttctcagaa 25

RESULT 173

ID US-08-736-723A-1 STANDARD; DNA; UNC; 1375 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08736723A
CC Sequence 1, Application US/08736723A
CC Patent No. 5869235
CC GENERAL INFORMATION:
CC APPLICANT: SUDA, NARUSHI
CC APPLICANT: ITOGA, YUTAKA
CC TITLE OF INVENTION: GENE OF THE HOP LATENT VIRUS AND METHODS
CC TITLE OF INVENTION: FOR DETECTING THE SAME
CC NUMBER OF SEQUENCES: 7
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBION, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
CC STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CC CITY: ARLINGTON
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/736,723A
CC FILING DATE: 25-OCT-1996
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 7-302297
CC FILING DATE: 27-OCT-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 7-352285
CC FILING DATE: 28-DEC-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: OBION, NORMAN F.
CC REGISTRATION NUMBER: 24,618
CC REFERENCE/DOCKET NUMBER: 2589-042-0
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-413-3000
CC TELEFAX: 703-413-2220
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1375 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
SQ SEQUENCE 1375 BP; 363 A; 282 C; 382 G; 348 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1375;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 102 TTCTCAGAA 110
|||||
Cp 33 ttctcagaa 25

RESULT 174
ID PCT-US96-03940-3 STANDARD; DNA; UNC; 1407 BP.
AC xxxxxx
DT
DE Sequence 3, Application PC/TUS9603940
CC Sequence 3, Application PC/TUS9603940
CC GENERAL INFORMATION:
CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC TITLE OF INVENTION: REST Protein and DNA
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Dechert, Price & Rhoads
CC STREET: P.O. Box 5218

CC CITY: Princeton
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 08543-5218
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
CC COMPUTER: IBM-compatible
CC OPERATING SYSTEM: DOS 5.0
CC SOFTWARE: WordPerfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/03940
CC FILING DATE: March 23, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Allen Bloom
CC REGISTRATION NUMBER: 29,135
CC REFERENCE/DOCKET NUMBER: 317743-101
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214
CC TELEFAX: (609) 520-3259
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1407 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: no
CC ANTI-SENSE: no
CC ORGANISM: Human
CC CELL LINE: HeLa
CC IMMEDIATE SOURCE:
CC LIBRARY: CDNA
CC PUBLICATION INFORMATION:
CC AUTHORS: Chong, Jayhong A., Tapia-Ramirez Jos, Toledo-
CC AUTHORS: Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
CC AUTHORS: Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
CC TITLE: REST: A Mammalian Silencer Protein that Restricts
CC JOURNAL: Cell
CC VOLUME: 80
CC ISSUE:
CC PAGES: March 24, 1995
CC RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 1407
CC SEQUENCE 1407 BP; 515 A; 276 C; 306 G; 310 T; 0 OTHER.
CC
CC Query Match 100.0%; Score 9; DB 4; Length 1407;
CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Db 724 TTCACGAA 732
CC |||||
CC Cp 33 ttctcagaa 25
CC
CC RESULT 175
CC ID US-08-252-995D-1 STANDARD; DNA; UNC; 1453 BP.
CC AC xxxxxx
CC DE Sequence 1, Application US/08252995D
CC CC Sequence 1, Application US/08252995D
CC CC Patent No. 5650501
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Dennis, James W
CC CC APPLICANT: Heffernan, Mike
CC CC APPLICANT: Fode, Carol
CC CC TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
CC CC NUMBER OF SEQUENCES: 14
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: BERSKIN & PARR
CC CC STREET: 40 King Street West

CC CITY: Toronto
CC STATE: Ontario
CC COUNTRY: Canada
CC ZIP: M5H 3Y2
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/252,995D
CC FILING DATE: 02-JUN-1994
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kurzydys, Linda M
CC REGISTRATION NUMBER: 34,971
CC REFERENCE/DOCKET NUMBER: 3153-96
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 364-7311
CC TELEFAX: (416) 361-1398
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1453 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC ORIGINAL SOURCE:
CC ORGANISM: Mus musculus
CC DEVELOPMENTAL STAGE: Lymphoid CDNA library
CC IMMEDIATE SOURCE:
CC LIBRARY: Murine Lymphoid
CC CLONE: WGA-Resistant Chop Clones
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 206..1453
CC FEATURE:
CC NAME/KEY: 5'UTR
CC LOCATION: 1..205
CC SEQUENCE 1453 BP; 426 A; 315 C; 357 G; 355 T; 0 OTHER.
CC
CC Query Match 100.0%; Score 9; DB 1; Length 1453;
CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC Db 527 TTCACGAA 535
CC |||||
CC Cp 33 ttctcagaa 25
CC
CC RESULT 176
CC ID PCT-US96-03940-27 STANDARD; DNA; UNC; 1461 BP.
CC AC xxxxxx
CC DE Sequence 27, Application PC/TUS9603940
CC CC Sequence 27, Application PC/TUS9603940
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Mandel, Gail, Chong, Jayhong A.
CC CC NUMBER OF SEQUENCES: 29
CC CC TITLE OF INVENTION: REST Protein and DNA
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Dechert Price & Rhoads
CC CC STREET: P.O. Box 5218
CC CC CITY: Princeton
CC CC STATE: New Jersey
CC CC COUNTRY: USA
CC CC ZIP: 08543-5218
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
CC CC COMPUTER: IBM-compatible
CC CC OPERATING SYSTEM: DOS 5.0
CC CC SOFTWARE: WordPerfect
CC CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US96/03940
CC FILING DATE: March 23, 1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Allen Bloom
CC REGISTRATION NUMBER: 29,135
CC REFERENCE/DOCKET NUMBER: 317743-101
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (609) 520-3214
CC TELEFAX: (609) 520-3259
CC INFORMATION FOR SEQ ID NO: 27:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1461 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: no
CC ANTI-SENSE: no
CC ORIGINAL SOURCE:
CC ORGANISM: Human
CC CELL LINE: HeLa
CC IMMEDIATE SOURCE:
CC LIBRARY: cDNA
CC PUBLICATION INFORMATION:
CC AUTHORS: Chong, Jayhong A., Tapia-Ramirez Jos, Toledo-
CC AUTHORS: Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
CC AUTHORS: Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
CC TITLE: REST: A Mammalian Silencer Protein that Restricts
CC TITLE: Sodium Channel Gene Expression to Neurons
CC JOURNAL: Cell
CC VOLUME: 80
CC ISSUE:
CC PAGES:
CC DATE: March 24, 1995
CC RELEVANT RESIDUES IN SEQ ID NO: 26: FROM 1 TO 1461
SQ SEQUENCE 1461 BP; 528 A; 288 C; 326 G; 319 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 1461;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 819 TTCTCAGAA 827
Cp 33 ttctcagaa 25
RESULT 177
ID US-07-768-286B-5 STANDARD; DNA; UNC; 1482 BP.
AC xxxxxx
DE Sequence 5, Application US/07768286B
CC Sequence 5, Application US/07768286B
CC Patent No. 5444153
CC GENERAL INFORMATION:
CC APPLICANT: GOSS, Neil H.
CC APPLICANT: RICHARDSON, Michael A.
CC TITLE OF INVENTION: VARIANTS OF PAI-2
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/768,286B

CC FILING DATE: 19911011
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/AU90/00603
CC FILING DATE: 20-DEC-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 16786/157 CHAC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 836-9300
CC TELEFAX: (703) 683-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1482 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC CELL TYPE: Monocyte
CC CELL LINE: U937
CC IMMEDIATE SOURCE:
CC CLONE: BTA 1922
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 22..1170
CC IDENTIFICATION METHOD: experimental
CC OTHER INFORMATION: /function= "Product binds to
CC OTHER INFORMATION: /product= "PAL-2 variant, protease sensitive site
CC OTHER INFORMATION: /removed
CC OTHER INFORMATION: /evidence= EXPERIMENTAL
CC OTHER INFORMATION: /note= "Codes for human plasminogen activator
CC OTHER INFORMATION: inhibitor type 2 protein in which amino acids 66
CC OTHER INFORMATION: to 98 inclusive have been deleted."
SQ SEQUENCE 1482 BP; 439 A; 307 C; 320 G; 416 T; 0 OTHER.
Query Match 100.0%; Score 9; DB 1; Length 1482;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 890 TTCTCAGAA 898
Qy 25 ttctcagaa 33
RESULT 178
ID PCT-US95-15646-1 STANDARD; DNA; UNC; 1501 BP.
AC xxxxxx
DE Sequence 1, Application PC/TUS9515646
CC Sequence 1, Application PC/TUS9515646
CC GENERAL INFORMATION:
CC APPLICANT: Synaptic Pharmaceutical Corporation
CC TITLE OF INVENTION: METHODS OF MODIFYING FEEDING BEHAVIOR, COMPOUNDS
CC TITLE OF INVENTION: USEFUL IN SUCH METHODS, AND DNA ENCODING A HYPOTHALAMIC
CC TITLE OF INVENTION: Y/PEPTIDE Y RECEPTOR (Y5) AND USES THEREOF
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham LLP
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: United States of America
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC * CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/15646
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28,678
CC REFERENCE/DOCKET NUMBER: 1795/46166-A-PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1501 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 61..1432
CC SEQUENCE 1501 BP; 400 A; 360 C; 313 G; 428 T; 0 OTHER.
S0
Query Match 100.0%; Score 9; DB 4; Length 1501;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 10 TTCGAGAA 18
QY 25 ttctgagaa 33
RESULT 179
ID US-08-349-025-1 STANDARD; DNA; UNC; 1501 BP.
AC xxxxxx
DE Sequence 1, Application US/08349025
CC Sequence 1, Application US/08349025
CC Patent No. 5602024
CC GENERAL INFORMATION:
CC APPLICANT: Gerald, Christophe P.G.
CC APPLICANT: Walker, Mary W.
CC APPLICANT: Branchek, Theresa
CC APPLICANT: Weinschank, Richard L.
CC TITLE OF INVENTION: DNA ENCODING A HYPOTHALAMIC ATYPICAL
CC TITLE OF INVENTION: NEUROPEPTIDE Y/PEPTIDE YX RECEPTOR (Y5) AND USES THEREOF
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: United States of America
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/349,025
CC FILING DATE:
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28,678
CC REFERENCE/DOCKET NUMBER: 1795/46166
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400

CC TELEFAX: (212) 391-0525
CC TELEX: 423523 COOP UI
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1501 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 61..1432
CC SEQUENCE 1501 BP; 400 A; 360 C; 313 G; 428 T; 0 OTHER.
S0
Query Match 100.0%; Score 9; DB 1; Length 1501;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 10 TTCGAGAA 18
QY 25 ttctgagaa 33
RESULT 180
ID PCT-US92-07103-1 STANDARD; DNA; UNC; 1510 BP.
AC xxxxxx
DE Sequence 1, Application PC/TUS9207103
CC Sequence 1, Application PC/TUS9207103
CC GENERAL INFORMATION:
CC APPLICANT: Massachusetts Institute of Technology; Forsyth Dental
CC APPLICANT: Center and Ohio State University
CC TITLE OF INVENTION: NOVEL STRAIN OF CHLAMYDIA
CC NUMBER OF SEQUENCES: 1
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
CC STREET: TWO MILLITIA DRIVE
CC CITY: LEXINGTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/07103
CC FILING DATE: 19920824
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/748,998
CC FILING DATE: 23-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Brook, David E.
CC REGISTRATION NUMBER: 22,592
CC REFERENCE/DOCKET NUMBER: MIT-5643A PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 861-6240
CC TELEFAX: (617) 861-9540
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1510 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: rRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Chlamydia sp.

CC STRAIN: novel
CC POSITION IN GENOME:
CC CHROMOSOME/SEGMENT: 165
CC PUBLICATION INFORMATION:
CC AUTHORS: Stills Jr., Harold F.
CC TITLE: Isolation of an Intracellular Bacterium
CC JOURNAL: Infect. Immun.
CC VOLUME: 59
CC ISSUE: 9
CC DATE: August 24-1991
SQ SEQUENCE 1510 BP; 369 A; 313 C; 450 G; 0 T; 378 OTHER.
Query Match 100.0%; Score 9; DB 4; Length 1510;
Best Local Similarity 66.7%; Pred. No. 3.96e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 UUCUGAGAA 9
:::|||||
QY 25 ttctgagaa 33

RESULT 181
ID US-07-768-286B-3 STANDARD; DNA; UNC; 1512 BP.
AC xxxxxx
DE Sequence 3, Application US/07768286B
CC Sequence 3, Application US/07768286B
CC Patent No. 544153
CC GENERAL INFORMATION:
CC APPLICANT: GOSS, Neil H.
CC APPLICANT: RICHARDSON, Michael A.
CC TITLE OF INVENTION: VARIANTS OF PAI-2
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 1800 Diagonal Road, Suite 500
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22313-0299
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/768,286B
CC FILING DATE: 19911011
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/AU90/00603
CC FILING DATE: 20-DEC-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 16786/157 CHAC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703)836-9300
CC TELEFAX: (703)883-4109
CC TELEX: 899149
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1512 base pairs
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC CELL TYPE: Monocyte
CC CELL LINE: U937

CC IMMEDIATE SOURCE:
CC CLONE: BTA 1916
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 22..1200
CC IDENTIFICATION METHOD: experimental
CC OTHER INFORMATION: /function- "product binds to
CC OTHER INFORMATION: urokinase, tissue plasminogen activator"
CC OTHER INFORMATION: /product- "PAI-2 variant, protease sensitive site
CC OTHER INFORMATION: removed"
CC OTHER INFORMATION: /evidence- EXPERIMENTAL
CC OTHER INFORMATION: /note- "Codes for human plasminogen activator
CC OTHER INFORMATION: inhibitor type 2 protein in which amino acids 74
CC OTHER INFORMATION: to 96 inclusive have been deleted."
SQ SEQUENCE 1512 BP; 449 A; 318 C; 324 G; 421 T; 0 OTHER.

Db 920 TTCTGAGAA 928
|||||||
QY 25 ttctgagaa 33

RESULT 182
ID US-08-466-337A-17 STANDARD; DNA; UNC; 1550 BP.
AC xxxxxx
DE Sequence 17, Application US/08466337A
CC Sequence 17, Application US/08466337A
CC Patent No. 5830756
CC GENERAL INFORMATION:
CC APPLICANT: Haskill, John S.
CC APPLICANT: Baldwin Jr., Albert S.
CC TITLE OF INVENTION: Inhibitor of NF-kB Transcriptional
CC TITLE OF INVENTION: Activator and uses thereof
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
CC STREET: 6300 Sears Tower/ 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: United States
CC ZIP: 60606-6402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/466,337A
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Pochopien, Donald J.
CC REGISTRATION NUMBER: 32,167
CC REFERENCE/DOCKET NUMBER: 0899.008/33518
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1550 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 95..1045

SO SEQUENCE 1550 BP; 380 A; 402 C; 416 G; 352 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1550;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1261 TTCTGAGAA 1269
0Y 25 ttctgagaa 33

RESULT 183
ID US-08-475-359-17 STANDARD; DNA; UNC; 1550 BP.
AC xxxxxx

Sequence 17, Application US/08475359
Sequence 17, Application US/08475359
Patent No. 5846714
GENERAL INFORMATION:
APPLICANT: Haskell, John S.
APPLICANT: Baldwin Jr., Albert S.
TITLE OF INVENTION: Inhibitor of NF- κ B transcriptional
ACTIVATOR AND USES THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower/ 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,359
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pochopien, Donald J.
REGISTRATION NUMBER: 32,167
REFERENCE/DOCKET NUMBER: 0899.004.33514
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-0448
TELEFAX: 312/474-6300
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 95..1045
SEQUENCE 1550 BP; 380 A; 402 C; 416 G; 352 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1550;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1261 TTCTGAGAA 1269
0Y 25 ttctgagaa 33

RESULT 184
ID US-08-252-995D-5 STANDARD; DNA; UNC; 1600 BP.
AC xxxxxx

DE Sequence 5, Application US/08252995D
CC Sequence 5, Application US/08252995D
CC Patent No. 5650501
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,995D
FILING DATE: 02-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
DEVELOPMENTAL STAGE: Lymphoid cDNA library
IMMEDIATE SOURCE:
LIBRARY: Murine Lymphoid
CLONE: MCA-resistant chop clones
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..205
FEATURE:
NAME/KEY: CDS
LOCATION: 206..1597
SEQUENCE 1600 BP; 482 A; 341 C; 379 G; 398 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1600;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 527 TTCTGAGAA 535
Cp 33 ttctcagaa 25

RESULT 185
ID US-07-768-286B-1 STANDARD; DNA; UNC; 1610 BP.
AC xxxxxx

Sequence 1, Application US/07768286B
Sequence 1, Application US/07768286B
Patent No. 5444153
GENERAL INFORMATION:
APPLICANT: Goss, Neil H.
APPLICANT: RICHARDSON, Michael A.

```

CC      TITLE OF INVENTION:  VARIANTS OF PAI-2
CC      NUMBER OF SEQUENCES:  21
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE:  Foley & Lardner
CC      STREET:  1800 Diagonal Road, Suite 500
CC      CITY:  Alexandria
CC      STATE:  Virginia
CC      COUNTRY:  USA
CC      ZIP:  22313-0299
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE:  Floppy disk
CC      COMPUTER:  IBM PC compatible
CC      OPERATING SYSTEM:  PC-DOS/MS-DOS
CC      SOFTWARE:  Patentin Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER:  US/07/768,286B
CC      FILING DATE:  19911011
CC      CLASSIFICATION:  514
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER:  PCT/AU90/00603
CC      FILING DATE:  20-DEC-1990
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME:  BENT, Stephen A.
CC      REGISTRATION NUMBER:  29,768
CC      REFERENCE/DOCKET NUMBER:  16786/157 CHAC
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE:  (703)683-9300
CC      TELEFAX:  (703)683-4109
CC      TELEX:  899149
CC      INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH:  1610 base pairs
CC      TYPE:  NUCLEIC ACID
CC      STRANDEDNESS:  double
CC      TOPOLOGY:  linear
CC      MOLECULE TYPE:  cDNA to mRNA
CC      HYPOTHETICAL:  NO
CC      ANTI-SENSE:  NO
CC      ORIGINAL SOURCE:
CC      ORGANISM:  Homo sapiens
CC      CELL TYPE:  Monocyte
CC      IMMEDIATE SOURCE:
CC      CLONE:  BTA 1445
CC      POSITION IN GENOME:
CC      CHROMOSOME/SEGMENT:  18
CC      MAP POSITION:  18q21-q23
CC      FEATURE:
CC      NAME/KEY:  CDS
CC      LOCATION:  49..1296
CC      OTHER INFORMATION:  /product= "human plasminogen
CC      OTHER INFORMATION:  activator inhibitor type 2 protein"
S0      SEQUENCE 1610 BP:  485 A; 342 C; 342 G; 441 T; 0 OTHER.

Db      1016 TTCTGAGAA 1024
QY      25 ttctgagaa 33

RESULT  186
ID      US-08-399-561-4 STANDARD; DNA; UNC; 1614 BP.
AC      xxxxxx
DT
DE      Sequence 4, Application US/08399561
CC      Sequence 4, Application US/08399561
CC      Patent No. 5792903
CC      GENERAL INFORMATION:
CC      APPLICANT:  Hirschberg, Joseph
CC      APPLICANT:  Cunningham Jr., Francis X.

```

```

CC      APPLICANT: Gantt, Elisabeth
CC      TITLE OF INVENTION: Tissue cyclase gene
CC      NUMBER OF SEQUENCES: 5
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Kohn & Associates
CC      STREET: 30500 No. 5792903thwestern Highway, Suite 410
CC      CITY: Farmington Hills
CC      STATE: MI
CC      COUNTRY: US
CC      ZIP: 48334
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/399,561
CC      FILING DATE:
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Kohn, Kenneth I.
CC      REGISTRATION NUMBER: 30,955
CC      REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 810-539-5050
CC      TELEFAX: 810-539-5050
CC      INFORMATION FOR SEQ ID NO: 4:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1614 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: cDNA to mRNA
CC      ORIGINAL SOURCE:
CC      ORGANISM: Nicotiana tabacum
CC      STRAIN: cv. xanthill
CC      TISSUE TYPE: leaf
CC      SEQUENCE 1614 BP; 447 A; 288 C; 370 G; 509 T; 0 OTHER.
CC
CC      Query Match      100.0%; Score 9; DB 3; length 1614;
CC      Best Local Similarity 100.0%; Fred. No. 3.96e+02;
CC      Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
CC
CC      Db      1224 TTCTGAGAA 1232
CC      QY      25 ttctgagaa 33
CC
CC      RESULT 187
CC      ID      US-08-709-923-2 STANDARD; DNA; UNC; 1625 BP.
CC      AC      xxxxxx
CC      DT
CC      DE      Sequence 2, Application US/08709923
CC      SE      Sequence 2, Application US/08709923
CC      PAT      Patent No. 5831055
CC      GENERAL INFORMATION:
CC      APPLICANT: BIENKOWSKI, MICHAEL J.
CC      TITLE OF INVENTION: NOVEL KIDNEY ATP-DEPENDENT POTASSIUM
CC      TITLE OF INVENTION: CHANNELS
CC      NUMBER OF SEQUENCES: 8
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pharmacia & Upjohn Company
CC      STREET: 301 Henrietta street
CC      CITY: Kalamazoo
CC      STATE: MI
CC      COUNTRY: USA
CC      ZIP: 49001
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:

```


CC APPLICATION NUMBER: US/08/709,923
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: WOOLTON, Thomas A.
CC REGISTRATION NUMBER: 35,004
CC REFERENCE/DOCKET NUMBER: 6001.N CP
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 616-833-7914
CC TELEFAX: 616-833-6897
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1625 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA to mRNA
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: N-terminal
CC SEQUENCE 1625 BP; 444 A; 399 C; 367 G; 415 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1625;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1111 TTCTGAGAA 1119
1111111111
0Y 25 ttctgagaa 33

RESULT 188
ID US-08-460-895-1 STANDARD; DNA; UNC; 1634 BP.
AC xxxxxx
DE Sequence 1, Application US/08460895
CC Sequence 1, Application US/08460895
CC Patent No. 5849508
CC GENERAL INFORMATION:
CC APPLICANT: BRECHOT, Christian
CC APPLICANT: WANG, Jian
CC APPLICANT: CHENIVESSE, Xavier
CC APPLICANT: HENGLEIN, Berthold
CC APPLICANT: ZINDY, Fr d rigne
CC TITLE OF INVENTION: NEW HUMAN CYCLIN A COMPOSITIONS AND A PROCESS FOR THEIR PRO
CC NUMBER OF SEQUENCES: 1
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Larson and Taylor
CC STREET: 727 Twenty-Third Street, South
CC CITY: Arlington
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette-5.25 inch, 500 Kb storage
CC COMPUTER: IBM PC/XT/AT or compatibles
CC OPERATING SYSTEM: MS-DOS version 3.0 or above
CC SOFTWARE: WORD PERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/460,895
CC FILING DATE: 05-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/650, 805
CC FILING DATE: 06-FEB-1991
CC APPLICATION NUMBER: FR9001596
CC FILING DATE: 12-FEB-1990
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1634 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: CDNA
CC ORIGINAL SOURCE:
CC ORGANISM: HOMO SAPIENS
CC FEATURE:
CC NAME/KEY: Human cyclin A
CC LOCATION: coding sequence from base 97 to base 1392,
CC LOCATION: coding for a protein of 432 amino acids.
CC SEQUENCE 1634 BP; 475 A; 354 C; 376 G; 429 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1634;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 968 TTCTGAGAA 976
1111111111
0Y 25 ttctgagaa 33

RESULT 189
ID 5198359-9 STANDARD; DNA; UNC; 1791 BP.
AC xxxxxx
DE 01-JAN-1900
DE Patent No. 5198359.
CC Patent No. 5198359
CC APPLICANT: TANIGUCHI, TADATSUGU; JATAKEYAMA, MASANORI;
CC MINAMOTO, SEIJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI;
CC TSUDO, MITSURU; KARASUYAMA, HAJIME
CC TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
CC NUMBER OF SEQUENCES: 9
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/487,059
CC FILING DATE: 05-MAR-1990
CC SEQ ID NO: 9:
CC LENGTH: 1653
CC SEQUENCE 1791 BP; 316 A; 560 C; 455 G; 322 T; 138 other;

Query Match 100.0%; Score 9; DB 5; Length 1653;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 279 TTCTGAGAA 287
1111111111
Cp 33 ttctgagaa 25

RESULT 190
ID 5449756-9 STANDARD; DNA; UNC; 1794 BP.
AC xxxxxx
DE 01-JAN-1900
DE Patent No. 5449756.
CC Patent No. 5449756
CC APPLICANT: TANIGUCHI, TADATSUGU; HATAKEYAMA, MASANORI; MINAMOTO,
CC SEJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI; TSUDO,
CC MITSURU; KARASUYAMA, HAJIME
CC TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
CC NUMBER OF SEQUENCES: 12
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/88,592
CC FILING DATE: 9-JUL-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 865,155
CC FILING DATE: 08-APR-1992
CC APPLICATION NUMBER: 487,059
CC FILING DATE: 05-MAR-1990
CC SEQ ID NO: 9:
CC LENGTH: 1656
CC SEQUENCE 1794 BP; 317 A; 559 C; 456 G; 324 T; 138 other;

Query Match 100.0%; Score 9; DB 5; Length 1656;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 282 TTCTGAGAA 290

CP 33 ttctcagaa 25

RESULT 191
ID 5198359-5 STANDARD; DNA; UNC: 1794 BP.

AC xxxxxx
DT 01-JUN-1900
DE Patent No. 5198359.

CC APPLICANT: TANIGUCHI, TADATSUGU, JATAKEYAMA, MASANORI;
CC MINAMOTO, SEIJIRO; KONO, TAKESHI; DOI, TAKESHI; MIYASAKA, MASAYUKI;
CC TSUDO, MITSURU; KARASUYAMA, HAJIME

CC TITLE OF INVENTION: RECOMBINANT PROTEIN RECEPTOR FOR IL-2
CC NUMBER OF SEQUENCES: 9
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/487,059
CC FILING DATE: 05-MAR-1990

CC SEQ ID NO: 5:
CC LENGTH: 1656
SQ Sequence 1794 BP: 317 A; 560 C; 456 G; 323 T; 138 other:

Query Match 100.0%; Score 9; DB 5; Length 1656;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 282 TTCTCAGAA 290
CP 33 ttctcagaa 25

RESULT 192
ID US-08-476-008-4 STANDARD; DNA; UNC: 1673 BP.

DE Sequence 4, Application US/08476008
CC Sequence 4, Application US/08476008
CC Patent No. 5627061

CC GENERAL INFORMATION:

CC APPLICANT: Barry, Gerard F.

CC APPLICANT: Kishore, Ganesh M.

CC APPLICANT: Padgett, Stephen R.

CC APPLICANT: Stallings, William C.

CC TITLE OF INVENTION: Glyphosate Tolerant

CC TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-phosphate Synthases

CC NUMBER OF SEQUENCES: 69

CC CORRESPONDENCE ADDRESSES:

CC ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B44F

CC STREET: 700 Chesterfield Village Parkway

CC CITY: St. Louis

CC STATE: Missouri

CC COUNTRY: USA

CC ZIP: 63198

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/476,008

CC FILING DATE: 07-JUN-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/306,063

CC FILING DATE: 13-SEP-1994

CC APPLICATION NUMBER: US 07/749,611

CC FILING DATE: 28-AUG-1991

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/576,537

CC FILING DATE: 31-AUG-1990

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Hoerner Jr., Dennis R.
CC REGISTRATION NUMBER: 30,914
CC REFERENCE/DOCKET NUMBER: 38-21(10660)A
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (314)537-6099

CC TELEFAX: (314)537-6047

CC INFORMATION FOR SEQ ID NO: 4:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1673 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC FEATURE:

CC NAME/KEY: CDS

CC LOCATION: 86..1432

SQ SEQUENCE 1673 BP: 354 A; 517 C; 485 G; 317 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1673;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1653 TTCTGAGAA 1661
OY 25 ttctgagaa 33

RESULT 193
ID US-08-833-485-4 STANDARD; DNA; UNC: 1673 BP.

DE Sequence 4, Application US/08833485
CC Sequence 4, Application US/08833485
CC Patent No. 5804425

CC GENERAL INFORMATION:

CC APPLICANT: Barry, Gerard F.

CC APPLICANT: Kishore, Ganesh M.

CC APPLICANT: Padgett, Stephen R.

CC APPLICANT: Stallings, William C.

CC TITLE OF INVENTION: Glyphosate Tolerant

CC TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-phosphate Synthases

CC NUMBER OF SEQUENCES: 69

CC CORRESPONDENCE ADDRESSES:

CC ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B44F

CC STREET: 700 Chesterfield Village Parkway

CC CITY: St. Louis

CC STATE: Missouri

CC COUNTRY: USA

CC ZIP: 63198

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/833,485

CC FILING DATE: 07-APR-1997

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/306,063

CC FILING DATE: 13-SEP-1994

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/749,611

CC FILING DATE: 28-AUG-1991

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/576,537

CC FILING DATE: 31-AUG-1990

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Hoerner Jr., Dennis R.

CC REGISTRATION NUMBER: 30,914

CC REFERENCE/DOCKET NUMBER: 38-21(15117)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314)737-6099
CC TELEFAX: (314)737-6047
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1673 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 86..1432
SQ SEQUENCE 1673 BP; 354 A; 517 C; 485 G; 317 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1673;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1653 TTCTGAGAA 1661
|||
OY 25 tctctgagaa 33

RESULT 194
ID US-08-306-063-4 STANDARD; DNA; UNC; 1673 BP.
AC xxxxxx
DT Sequence 4, Application US/08306063
DE Sequence 4, Application US/08306063
CC Patent No. 5633435
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Padgett, Stephen R.
CC APPLICANT: Stallings, William C.
CC TITLE OF INVENTION: Glyophosate Tolerant
CC TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
CC NUMBER OF SEQUENCES: 69
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
CC STREET: 700 Chesterfield Village Parkway
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/306,063
CC FILING DATE: 13-SEP-1994
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/749,611
CC FILING DATE: 28-AUG-1991
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/576,537
CC FILING DATE: 31-AUG-1990
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hoerner Jr., Dennis R.
CC REGISTRATION NUMBER: 30,914
CC REFERENCE/DOCKET NUMBER: 38-21(10660)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314)537-6099
CC TELEFAX: (314)537-6047
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1673 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 86..1432
SQ SEQUENCE 1673 BP; 354 A; 517 C; 485 G; 317 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1673;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1653 TTCTGAGAA 1661
|||
OY 25 tctctgagaa 33

RESULT 195
ID PCT-US91-06148A-4 STANDARD; DNA; UNC; 1673 BP.
AC xxxxxx
DT Sequence 4, Application PC/TUS9106148A
DE Sequence 4, Application PC/TUS9106148A
CC GENERAL INFORMATION:
CC APPLICANT: Barry, Gerard F.
CC APPLICANT: Kishore, Ganesh M.
CC APPLICANT: Padgett, Stephen R.
CC TITLE OF INVENTION: Glyophosate Tolerant
CC TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
CC NUMBER OF SEQUENCES: 36
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Dennis R. Hoerner, Jr., Monsanto Co. BBAF
CC STREET: 700 Chesterfield Village Parkway
CC CITY: St. Louis
CC STATE: Missouri
CC COUNTRY: USA
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US91/06148A
CC FILING DATE: 19910828
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US/07/749,611
CC FILING DATE: 28-AUG-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/576,537
CC FILING DATE: 31-AUG-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hoerner Jr., Dennis R.
CC REGISTRATION NUMBER: 30,914
CC REFERENCE/DOCKET NUMBER: 38-21(10535)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314)537-6099
CC TELEFAX: (314)537-6047
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1673 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 86..1432
SQ SEQUENCE 1673 BP; 354 A; 517 C; 485 G; 317 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 4; Length 1673;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1653 TTCTGAGAA 1661
|||||
OY 25 ttctgagaa 33

RESULT 196
ID US-08-484-993B-40 STANDARD; DNA; UNC; 1701 BP.
AC xxxxxx
DT
DE Sequence 40, Application US/08484993B
CC Sequence 40, Application US/08484993B
CC Patent No. 5837497
CC GENERAL INFORMATION:
CC APPLICANT: Harris Ph.D., Jeffrey D.
CC APPLICANT: Hsu, Kuang T.
CC APPLICANT: Podolski, Joseph S.
CC TITLE OF INVENTION: Materials and Methods for Immunocotraception
CC NUMBER OF SEQUENCES: 59
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: United States of America
CC ZIP: 60606-6402
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/484,993B
CC FILING DATE: 09-NOV-1993
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/012,990
CC FILING DATE: 29-JAN-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/973,341
CC FILING DATE: 09-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clough, David W.
CC REGISTRATION NUMBER: 36,107
CC REFERENCE/DOCKET NUMBER: 31745
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6553
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 40:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1701 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..1698
CC SEQUENCE 1701 BP; 392 A; 470 C; 421 G; 418 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1701;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1491 TTCTGAGAA 1499
|||||
Cp 33 ttctgagaa 25

RESULT 197
ID US-08-247-809A-5 STANDARD; DNA; UNC; 1722 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08247809A
CC Sequence 5, Application US/08247809A
CC Patent No. 5569823
CC GENERAL INFORMATION:
CC APPLICANT: Peter H. Schreier; Klaus Stenzel; Gunter Adam;
CC APPLICANT: Edgar Maiss
CC TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SPRING HORN KRAMER & WOODS
CC STREET: 660 White Plains Road
CC CITY: Tarrytown
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10591-5144
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
CC MEDIUM TYPE: storage
CC COMPUTER: NEC Powermate 1 Plus
CC OPERATING SYSTEM: DOS
CC SOFTWARE: Wordperfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/247,809A
CC FILING DATE: May 23, 1994
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: P 43 178 45.6 (Germany)
CC FILING DATE: May 28, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kurt G. Briscoe
CC REGISTRATION NUMBER: 33,141
CC REFERENCE/DOCKET NUMBER: Bayer 9049-KGB
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (914) 332-1700
CC TELEFAX: (914) 332-1844
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1722 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 475..1428
CC SEQUENCE 1722 BP; 580 A; 343 C; 338 G; 461 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 1; Length 1722;
Best Local Similarity 100.0%; Pred. No. 3.96e+02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 984 TTCTGAGAA 992
|||||
OY 25 ttctgagaa 33

RESULT 198
ID US-08-417-330A-17 STANDARD; DNA; UNC; 1732 BP.
AC xxxxxx
DT
DE Sequence 17, Application US/08417330A
CC Sequence 17, Application US/08417330A
CC Patent No. 5719057
CC GENERAL INFORMATION:
CC APPLICANT: HADINGHAM, KAREN
CC APPLICANT: LE BOURDELLES, BEATRICE
CC APPLICANT: WHITING, PAUL
CC APPLICANT: WINGROVE, PETER

CC TITLE OF INVENTION: STABLY TRANSPECTED CELL LINE EXPRESSING
CC TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLOVED
CC TITLE OF INVENTION: GABA-RECEPTOR SUBUNIT CDNA SEQUENCES
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
CC STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CC CITY: RAHWAY
CC STATE: NJ
CC COUNTRY: US
CC ZIP: 07065-0900
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/417,330A
CC FILING DATE: 05-APR-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC ATTORNEY/AGENT INFORMATION:
CC FILING DATE:
CC NAME: HAND, MARK
CC REGISTRATION NUMBER: 36,545
CC REFERENCE/DOCKET NUMBER: T11091A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 908-594-3905
CC TELEFAX: 908-594-4720
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 17:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1732 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC FEATURE:
CC NAME/KEY: Coding Sequence
CC LOCATION: 27...1385
CC OTHER INFORMATION:
SQ SEQUENCE 1732 BP: 508 A: 366 C: 354 G: 504 T: 0 OTHER.
CC Query Match 100.0%; Score 9; DB 2; Length 1732;
CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1477 TTCTCAGAA 1485
CP 33 ttctcagaa 25
RESULT 199
ID 5258283-8 STANDARD; DNA; UNC; 1894 BP.
AC xxxxxx
DE 01-JAN-1900
DE Patent No. 5258283.
CC APPLICANT: FRAZIER, MARVIN E.; MALAVIA, LOUIS P.; SAMUEL,
CC JAMES E.; BACA, OSWALD G.
CC TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF COXIELLA
CC BURNETTII IN BIOLOGICAL FLUIDS
CC NUMBER OF SEQUENCES: 17
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/425,856
CC FILING DATE: 23-OCT-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 927,779
CC FILING DATE: 05-NOV-1986
CC APPLICATION NUMBER: 795,207
CC FILING DATE: 05-NOV-1985
CC SEQ ID NO:8:

CC LENGTH: 1749
CC Sequence 1894 BP: 579 A: 335 C: 343 G: 492 T: 145 other;
SQ Query Match 100.0%; Score 9; DB 5; Length 1749;
CC Best Local Similarity 100.0%; Pred. No. 3.96e+02;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 807 TTCTCAGAA 815
CP 33 ttctcagaa 25
RESULT 200
ID US-08-453-848-14 STANDARD; DNA; UNC; 1757 BP.
AC xxxxxx
DE Sequence 14, Application US/08453848
DE Sequence 14, Application US/08453848
DE Patent No. 5858368
CC GENERAL INFORMATION:
CC APPLICANT: Smith, Gale Eugene
CC APPLICANT: Voliovitz, Franklin
CC APPLICANT: Wilkinson, Bethanie Eident
CC APPLICANT: Voznesensky, Andrei I.
CC APPLICANT: Hackett, Craig Stanway
CC TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
CC TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
CC NUMBER OF SEQUENCES: 31
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Patrea L. Pabst
CC STREET: 2800 One Atlantic Center
CC STREET: 1201 West Peachtree Street
CC CITY: Atlanta
CC STATE: GA
CC COUNTRY: USA
CC ZIP: 30309-3450
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/453,848
CC FILING DATE: 30-MAY-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/120,607
CC FILING DATE: 13-SEPT-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Pabst, Patrea L.
CC REGISTRATION NUMBER: 31,284
CC REFERENCE/DOCKET NUMBER: MGS101CIP
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (404)-873-8794
CC TELEFAX: (404)-873-8795
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1757 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Influenza virus
CC INDIVIDUAL ISOLATE: A/Shandong/9/93 rHA
CC FEATURE:
CC NAME/KEY: polyhedrin mRNA leader (partial)
CC LOCATION: 1 to 18
CC FEATURE:
CC NAME/KEY: coding region for AcNPV 61k protein signal

CC NAME/KEY: sequence
 CC LOCATION: 19 to 72
 CC FEATURE:
 CC NAME/KEY: SmaI restriction site
 CC LOCATION: 76 to 81
 CC FEATURE:
 CC NAME/KEY: coding region for mature rHA
 CC LOCATION: 73 to 1728
 CC FEATURE:
 CC NAME/KEY: KpnI restriction site
 CC LOCATION: 1735 to 1740
 CC FEATURE:
 CC NAME/KEY: BglII restriction site
 CC LOCATION: 1741 to 1746
 CC FEATURE:
 CC NAME/KEY: universal translation termination signal
 CC LOCATION: 1747 to 1757
 CC NAME/KEY: universal translation termination signal
 CC LOCATION: 1747 to 1757
 SQ SEQUENCE 1757 BP: 582 A; 352 C; 403 G; 420 T; 0 OTHER.

Query Match 100.0%; Score 9; DB 3; Length 1757;
 Best Local Similarity 100.0%; Pred. No. 3.96e+02;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1276 TTCTCAGAA 1284
 |||||
 Cp 33 ttctcagaa 25

Search completed: Thu Oct 28 00:34:43 1999
 Job time : 59 secs.

84	9	100.0	250.20	AA872080	ct12a11.s1 NCI_CGAP_GC	2.23e+03	157	9	100.0	300.15	C36735	C36735	Yuj1 Kohara unp	2.23e+03
85	9	100.0	250.16	AA640401	nt99e08.s1 NCI_CGAP_AI	2.23e+03	158	9	100.0	300.15	C34709	C34709	Yuj1 Kohara unp	2.23e+03
86	9	100.0	252.11	AA343598	Est49407 Gall Bladder	2.23e+03	159	9	100.0	300.15	C31427	C31427	Yuj1 Kohara unp	2.23e+03
87	9	100.0	252.20	AA905871	o18a08.s1 Soares_NFL	2.23e+03	160	9	100.0	300.15	C54752	C54752	Yuj1 Kohara unp	2.23e+03
88	9	100.0	254.17	AA743482	ny21h01.s1 NCI_CGAP_GC	2.23e+03	161	9	100.0	300.15	C56383	C56383	Yuj1 Kohara unp	2.23e+03
89	9	100.0	254.21	AT014786	ot79f02.s1 Soares_tota	2.23e+03	162	9	100.0	300.12	AA382476	C54752	Yuj1 Kohara unp	2.23e+03
90	9	100.0	255.14	AA527319	ng36b05.s1 NCI_CGAP_CO	2.23e+03	163	9	100.0	300.15	C54178	C54178	Yuj1 Kohara unp	2.23e+03
91	9	100.0	256.15	AA593154	nm06c07.s1 NCI_CGAP_PT	2.23e+03	164	9	100.0	300.14	AA516287	nt68h04.s1 NCI_CGAP_CO	2.23e+03	
92	9	100.0	257.13	AA429636	zw74h05.s1 Soares_test	2.23e+03	165	9	100.0	300.15	C33705	C33705	Yuj1 Kohara unp	2.23e+03
93	9	100.0	259.21	AA919105	os76g01.s1 NCI_CGAP_GC	2.23e+03	166	9	100.0	300.15	C53842	C33842	Yuj1 Kohara unp	2.23e+03
94	9	100.0	261.14	AA496492	Est65872 Jurkat T cell	2.23e+03	167	9	100.0	300.21	C94621	C34621	Yuj1 Kohara unp	2.23e+03
95	9	100.0	261.12	AA357180	EST97241 Testis I Homo	2.23e+03	168	9	100.0	301.12	AA385118	EST98723 Thyroid Homo	2.23e+03	
96	9	100.0	262.12	AA383812	Est97241 Testis I Homo	2.23e+03	169	9	100.0	301.11	AA333727	EST37844 Embryo, 9 wee	2.23e+03	
97	9	100.0	262.17	AA721968	zh17b04.s1 Soares_pine	2.23e+03	170	9	100.0	301.21	AA1019762	EST90004.r1 Soares_mous	2.23e+03	
98	9	100.0	263.19	R03503	pk08c03.r1 Kwabara MI	2.23e+03	171	9	100.0	302.12	AA381889	EST95202 Activated T-c	2.23e+03	
99	9	100.0	263.14	AA197928	mv10b04.r1 Guaywoodfor	2.23e+03	172	9	100.0	302.16	AA626881	z88h912.s1 Soares_test	2.23e+03	
100	9	100.0	264.8	AA525063	nh32g03.s1 NCI_CGAP_PT	2.23e+03	173	9	100.0	302.17	AA712997	33557 Lambda-PRU2 Arab	2.23e+03	
101	9	100.0	264.8	D33548	CEUK032E9R Yuj1 Kohara	2.23e+03	174	9	100.0	304.21	AA992859	SMEST0219 Schistosoma	2.23e+03	
102	9	100.0	264.8	D339705	HSC1Hb02 normalized 1	2.23e+03	175	9	100.0	304.21	AA992859	ct48h11.s1 Soares_test	2.23e+03	
103	9	100.0	264.10	AA266518	ms67a11.r1 Soares_mous	2.23e+03	176	9	100.0	305.11	AA313179	EST18492 Liver, subtra	2.23e+03	
104	9	100.0	264.15	AA587547	nd44a11.s1 NCI_CGAP_AR	2.23e+03	177	9	100.0	305.18	AA760552	vt78e11.r1 Strataene	2.23e+03	
105	9	100.0	265.10	AA280530	zt11f12.s1 NCI_CGAP_GC	2.23e+03	178	9	100.0	306.11	AA313174	EST18461 Liver, subtra	2.23e+03	
106	9	100.0	266.15	AA603628	np19g01.s1 NCI_CGAP_PT	2.23e+03	179	9	100.0	306.20	AA881995	VS33910.r1 Strataene	2.23e+03	
107	9	100.0	266.20	AA904030	oe74e11.s1 NCI_CGAP_Lu	2.23e+03	180	9	100.0	306.19	F11767	HSC31C111 normalized 1	2.23e+03	
108	9	100.0	266.10	AA241715	MB12S18C9r3 Brugia mal	2.23e+03	181	9	100.0	308.8	L25482	HNM21ES151 Clontech HL	2.23e+03	
109	9	100.0	270.10	AA197883	mv04b11.r1 Guaywoodfor	2.23e+03	182	9	100.0	308.19	F07471	HSC27X041 normalized 1	2.23e+03	
110	9	100.0	270.10	AA276895	va23h12.r1 Guaywoodfor	2.23e+03	183	9	100.0	308.19	T90079	Yd33906.s1 Soares_feta	2.23e+03	
111	9	100.0	271.16	AA642354	ns29e08.s1 NCI_CGAP_GC	2.23e+03	184	9	100.0	309.13	AA419174	zv34g05.r1 Soares_ovat	2.23e+03	
112	9	100.0	271.14	AA497742	v172c09.r1 Strataene	2.23e+03	185	9	100.0	310.10	AA256848	z818i10.r1 Soares_NHm	2.23e+03	
113	9	100.0	271.16	AA182421	zp56f02.r1 Strataene	2.23e+03	186	9	100.0	310.14	T88794	Yd33208.s1 Soares_feta	2.23e+03	
114	9	100.0	272.16	AA640402	nt99e05.s1 NCI_CGAP_AI	2.23e+03	187	9	100.0	310.14	AA502321	nt26g06.s1 NCI_CGAP_CO	2.23e+03	
115	9	100.0	273.21	AA990483	ua60h12.s1 Soares_2bhm	2.23e+03	188	9	100.0	311.9	AA146289	mt06e07.r1 Soares_mous	2.23e+03	
116	9	100.0	274.16	AA684949	EST10536 Rat PC-12 ce	2.23e+03	189	9	100.0	312.8	T32011	EST42478 Human uterus	2.23e+03	
117	9	100.0	275.13	AA422037	zv28b06.s1 Soares_ovat	2.23e+03	190	9	100.0	312.17	AA724581	a101h11.s1 Soares_NFL	2.23e+03	
118	9	100.0	275.11	AA302178	EST10242 Adipose tissu	2.23e+03	191	9	100.0	313.17	AA713645	nv80h11.s1 NCI_CGAP_BT	2.23e+03	
119	9	100.0	276.11	AA282274	zt12g05.r1 NCI_CGAP_GC	2.23e+03	192	9	100.0	313.21	A1006978	ua77f02.r1 Soares_ZNDM	2.23e+03	
120	9	100.0	277.19	F10065	HSC3AB042 normalized 1	2.23e+03	193	9	100.0	315.15	AA508995	MBAPFX6A0713 Brugia ma	2.23e+03	
121	9	100.0	278.21	AA954232	oo92b04.s1 NCI_CGAP_K1	2.23e+03	194	9	100.0	316.10	AA252619	z144d06.s1 NCI_CGAP_GC	2.23e+03	
122	9	100.0	279.21	AA934905	oc10f10.s1 NCI_CGAP_GC	2.23e+03	195	9	100.0	316.10	AA230352	MY99e01.r1 Guaywoodfor	2.23e+03	
123	9	100.0	279.18	AA764915	nz60b04.s1 NCI_CGAP_GC	2.23e+03	196	9	100.0	317.8	T12492	B079R Heart Homo sapie	2.23e+03	
124	9	100.0	280.21	AA953790	oo88a11.s1 NCI_CGAP_Lu	2.23e+03	197	9	100.0	317.17	AA728030	K4S1BXY54SL JH0935L-B	2.23e+03	
125	9	100.0	280.16	AA629000	zv33c09.s1 Soares_test	2.23e+03	198	9	100.0	317.16	AA626165	zt90d10.s1 Soares_test	2.23e+03	
126	9	100.0	280.15	AA589358	v164e10.s1 Strataene	2.23e+03	199	9	100.0	319.11	AA301326	EST14257 Testis tumor	2.23e+03	
127	9	100.0	282.11	AA302617	EST15792 Aorta endothe	2.23e+03	200	9	100.0	319.20	AA929463	vt30h07.r1 Barstead mo	2.23e+03	
128	9	100.0	283.21	AT012399	EST206750 Normalized r	2.23e+03								
129	9	100.0	283.16	AA652376	ns52h09.s1 NCI_CGAP_PT	2.23e+03								
130	9	100.0	283.10	AA296586	EST112537 Adrenal_glan	2.23e+03								
131	9	100.0	284.20	AA927636	om19g12.s1 Soares_NFL	2.23e+03								
132	9	100.0	284.8	T60003	yc01e11.r1 Strataene	2.23e+03								
133	9	100.0	285.20	AA907008	ok98c12.s1 NCI_CGAP_Lu	2.23e+03								
134	9	100.0	292.13	AA429624	zw74g10.s1 Soares_test	2.23e+03								
135	9	100.0	292.10	AA297030	EST112879 Embryo, 9 we	2.23e+03								
136	9	100.0	292.11	AA344440	EST150332 Gall Bladder	2.23e+03								
137	9	100.0	293.8	T30360	EST15394 Human Endothe	2.23e+03								
138	9	100.0	294.12	AA369164	EST180533 Placenta II H	2.23e+03								
139	9	100.0	295.16	AA627369	ng59e01.s1 NCI_CGAP_CO	2.23e+03								
140	9	100.0	295.21	AA960651	ub60c01.s1 Soares_mous	2.23e+03								
141	9	100.0	295.8	M62263	EST00332 Hippocampus,	2.23e+03								
142	9	100.0	296.8	Z39383	HSC18H012 Normalized 1	2.23e+03								
143	9	100.0	298.20	AA905891	o182f08.s1 Soares_NFL	2.23e+03								
144	9	100.0	298.9	AA161311	zq38b09.r1 Strataene	2.23e+03								
145	9	100.0	298.12	AA363329	EST73119 Ovary I Homo	2.23e+03								
146	9	100.0	298.9	AA168582	ms31h06.r1 Strataene	2.23e+03								
147	9	100.0	299.21	AA977887	og67g07.s1 NCI_CGAP_K1	2.23e+03								
148	9	100.0	299.16	AA639969	np12g12.s1 NCI_CGAP_PT	2.23e+03								
149	9	100.0	299.11	AA336176	EST40702 Epididymus Ho	2.23e+03								
150	9	100.0	300.15	C38297	C38297 Yuj1 Kohara unp	2.23e+03								
151	9	100.0	300.15	C36182	C36182 Yuj1 Kohara unp	2.23e+03								
152	9	100.0	300.15	C37813	C37813 Yuj1 Kohara unp	2.23e+03								
153	9	100.0	300.21	C95112	C95112 Plc back fat Su	2.23e+03								
154	9	100.0	300.15	C33644	C33644 Yuj1 Kohara unp	2.23e+03								
155	9	100.0	300.15	C30237	C30237 Yuj1 Kohara unp	2.23e+03								
156	9	100.0	300.15	C32017	C32017 Yuj1 Kohara unp	2.23e+03								

ALIGNMENTS

RESULT 1
 LOCUS AA749827 63 bp mRNA
 DEFINITION vt48a09.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
 IMAGE:1123864 5', mRNA sequence.

ACCESSION AA794827
 NID 92857782
 VERSION AA794827.1 GI:2857782
 KEYWORDS EST.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 63)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kuubara,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 14, 1998 this sequence version replaced gi:1797442.

CONTACT: Marra M/Mouse EST Project
 WashU-HMT Mouse EST Project
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:613200.
 Location/Qualifiers

FEATURES
 source

1. .63
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /note="Organ: embryo; Vector: pBluescribe (modified);
 Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
 mRNA prepared from 13,500 2-cell stage embryos. Primer:
 SalI(dT): 5'-CGGTCGACGCGCGCGGTTTGT-3'. CDNAS
 were cloned into the MluI/SalI sites of a modified
 pBluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 kb."
 /db_xref="taxon:10090"
 /clone="IMAGE:1123864"
 /clone_lib="Knowles Solter mouse 2 cell"
 /tissue_type="embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"
 /lab_host="DH10B"

BASE COUNT 20 a 9 c 14 g 20 t
 ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 63;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS 2 AA472837 70 bp mRNA EST 18-JUN-1997
 DEFINITION v58f06.t1 Knowles Solter mouse blastocyst B1 Mus musculus CDNA
 Clone IMAGE:804803 5' similar to TR:EI28396 EI28396 XP0C2 PROTEIN.
 ; mRNA sequence.

ACCESSION AA472837
 NID 92201064
 VERSION AA472837.1 GI:2201064
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 70)
 Author(s) Matra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The Mashu-HMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:692466.

FEATURES
 source
 Contact: Matra M/Mouse EST Project
 Mashu-HMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:61147
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 High quality sequence stop: 1.
 Location/Qualifiers

source

1. .70
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally from mRNA prepared
 from 800 blastocysts. Primer: SalI(dT):
 5'-CGGTCGACGCGCGCGGTTTGT-3'. CDNAS were
 cloned into the NotI/SalI sites of a pSPORT vector (Life
 Technologies). Two different size selections: B1 (larger
 inserts) and B3."
 /db_xref="taxon:10090"
 /clone="IMAGE:804803"
 /clone_lib="Knowles Solter mouse blastocyst B1"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"
 /lab_host="DH10B"

BASE COUNT 23 a 12 c 18 g 17 t
 ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 70;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS 3 AA974588 114 bp mRNA EST 07-JUL-1998
 DEFINITION op28h09.s1 Soares NFL_T_GBC_S1 Homo sapiens CDNA clone
 IMAGE:1578209 3' mRNA sequence.

ACCESSION AA974588
 NID 93149768
 VERSION AA974588.1 GI:3149768
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 114)
 Author(s) NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:692499.

FEATURES
 source
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 491 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 85.
 Location/Qualifiers

1. .114
 /organism="Homo sapiens"
 /note="Organ: pooled; Vector: p7T73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBH19W, testis NHT, and B-cell
 NCI-CCAP-GCB1) were mixed and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified CDNAS from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /map="2"
 /clone="IMAGE:1578209"

BASE COUNT 30 a 27 c 21 g 36 t
 ORIGIN /clone.lib="Soares_NFL_T_GRC_S1"
 /lab_host="DH10B"

Query Match 100.0%; Score 9; DB 21; Length 114;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 84 TTCTGAGAA 92
 |||||||
 25 ttctgagaa 33

RESULT 4
 LOCUS AA623551 119 bp mRNA EST 14-OCT-1997
 DEFINITION v072e08.r1 Knowles Solter mouse blastocyst B1 Mus musculus CDNA
 clone IMAGE:1037510 5', mRNA sequence.
 ACCESSION AA623551
 NID 92527427
 VERSION AA623551.1 GI:2527427
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 119)
 Maier, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMNI Mouse EST Project
 Unpublished (1996)
 On Apr 14, 1993 this sequence version replaced g1:693581.

TITLE
 JOURNAL
 COMMENT

CONTACT: Marra M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MG:577534
 Putative full length read
 vector to vector length is 389
 High quality sequence stop: 106.
 Location/Qualifiers
 1. 119
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /note="Organ: embryo; Vector: pSPORT; Site: 1; NotI:
 Site 2: SalI; Cloned unidirectionally from mRNA prepared
 from 800 blastocysts. Primer: SalI(dT):
 5'-CGGTGACCGTCGACCGTTTTTTTTT-3'. cDNAs were
 cloned into the NotI/SalI sites of a pSPORT vector (Life
 Technologies). Two different size selections: B1 (larger
 inserts) and B3."
 /db_xref="taxon:10090"
 /clone.lib="IMAGE:1037510"
 /clone.lib="Knowles Solter mouse blastocyst B1"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"

BASE COUNT 36 a 25 c 20 g 38 t
 ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 119;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 85 TTCTGAGAA 93
 |||||||
 Cp 33 ttctgagaa 25

RESULT 5
 LOCUS AA62562 124 bp mRNA EST 12-NOV-1997
 DEFINITION nt16g04.s1 NCI-CGAP_Ew1 Homo sapiens CDNA clone IMAGE:1168182
 similar to SW:RL21_HUMAN P46778 60S RIBOSOMAL PROTEIN L21. mRNA
 sequence.
 ACCESSION AA62562
 NID 92615511
 VERSION AA62562.1 GI:2615511
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 124)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Sep 12, 1996 this sequence version replaced g1:1393031.

TITLE
 JOURNAL
 COMMENT

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.llnl.gov/dbp/image/image.html

Trace considered overall poor quality
 Seq primer: -40m3 fwd. ET from AmerSham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 124
 /organism="Homo sapiens"
 /note="Vector: pAMP10; mRNA made from Ewing's sarcoma,
 cDNA made by oligo-dT priming. Non-directionally cloned.
 Size-selected on agarose gel, average insert size 600 bp.
 Reference: Krizman et al. (1996) Cancer Research
 56:5360-5383."
 /db_xref="taxon:9606"
 /clone.lib="IMAGE:1168182"
 /clone.lib="NCI-CGAP_Ew1"
 /tissue_type="Ewing's sarcoma"
 /lab_host="DH10B"

BASE COUNT 24 a 30 c 18 g 52 t
 ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 124;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 49 TTCTGAGAA 57
 |||||||
 Cp 33 ttctgagaa 25

RESULT 6
 LOCUS AA509777 130 bp mRNA EST 08-JUL-1997
 DEFINITION v052e10.r1 Soares mouse mammary gland NbMNG Mus musculus CDNA clone
 IMAGE:890634 5' similar to gb:L20450 Mus musculus DNA-binding
 protein mRNA, complete cds (MOUSE);, mRNA sequence.
 ACCESSION AA509777
 NID 92247631
 VERSION AA509777.1 GI:2247631

```

KEYWORDS      EST.
SOURCE        Mus musculus
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 130)
AUTHORS        Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
                Gessel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
                Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
                Thaising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
                Waterston,R.
TITLE         The WashU-HMI Mouse EST Project
JOURNAL       Unpublished (1996)
COMMENT       On Sep 12, 1996 this sequence version replaced gi:1394098.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lln.gov) for further information.
MGI:518594
Trace considered overall poor quality
Seq primer: -28m13 rev2 Er from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..130
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15',
TGTTACCAATCTGAAAGTGGACGCCCGAATGTTTTTTTTTTTTTTTTTTTT
T 3'] : double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73D vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
/db_xref="taxon:10090"
/clone_1="IMAGE:890634"
/clone_1lb="Soares mouse mammary gland NBMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
1ab_host="DH10B"

BASE COUNT      41 a      24 c      25 g      40 t
ORIGIN
Query Match      100.0%; Score 9; DB 14; Length 130;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      37 TTCCTCAGAA 45
        |||||||
Cp      33 ttctcagaa 25

RESULT      7
LOCUS      AA204121      140 bp      mRNA      EST      27-JAN-1997
DEFINITION mus7d04.r1 Soares mouse lymph node NBM1N Mus musculus cDNA clone
IMAGE:643395 5' similar to gb:X56548 M.musculus Np-b mRNA for
purine-nucleoside phosphorylase (MUSE);, mRNA sequence.
ACCESSION  AA204121
NID         g1800757
VERSION     AA204121.1 GI:1800757
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 140)
Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gaisel,S., Kucada,T., Lacy,M., Le,M., Martin,U., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterson,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Apr 14, 1993 this sequence version replaced gi:785690.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL : contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
MG1:395487
Trace considered overall poor quality
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..140
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: lymph node; Vector: pT7p3D-Pac (Pharmacia)
with a modified polylinker. Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'
3'] double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7p3 vector. RNA
provided by Dr. Bertrand Jordan. Library constructed and
normalized by Bento Soares and M.Fatima Bonaudo."
/db_xref="taxon:10090"
/map="17q21"
/clone="IMAGE:643495"
/clone_lib="Soares mouse lymph node NbMLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 44 a 28 c 39 t
ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 140;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 115 TTCTGAGAA 123
|||||||
Oy 25 ttctgagaa 33

RESULT 8
LOCUS AA628553 142 bp mRNA EST 16-OCT-1997
DEFINITION af39all.8.1 Soares total.fetus.NB2HF8.9w Homo sapiens cDNA clone
IMAGE:1034012.3 similar to WP:W02B12.9 CE03769 MITOCHONDRIAL RNA
SPICING MSRA LIKE PROTEIN ;, mRNA sequence.
ACCESSION AA628553
NID g2540940
VERSION AA628553.1 GI:2540940
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 142)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maria, M., Martin, D., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Teising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R. Washu-NCI human EST Project
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1393210.
COMMENT

FEATURES
Source Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 145

BASE COUNT 36 a 42 c 34 g 30 t
ORIGIN
 Query Match 100.0%; Score 9; DB 16; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 TTCTGAGAA 19
 1111111111
Qy 25 ttctgagaa 33

RESULT 9 AA229563 145 bp mRNA EST 20-AUG-1997
LOCUS nc15e08.s1 NCI-CGAP_Prl Homo sapiens CDNA clone IMAGE:1008230
DEFINITION similar to gb:L29008 SORBITOL DEHYDROGENASE (HUMAN);, mRNA
ACCESSION AA229563
NID 91851704
VERSION AA229563.1 GI:1851704
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 145)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394888.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
 M.D., Michael Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Kitzman, Ph.D.
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/db/nci/image/image.html

FEATURES
Source Insert Length: 852 Std Error: 0.00
 Seq primer: -41ml3 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 145

BASE COUNT 40 a 28 c 31 g 46 t
ORIGIN
 Query Match 100.0%; Score 9; DB 10; Length 145;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 98 TTCTGAGAA 106
 1111111111
Qy 25 ttctgagaa 33

RESULT 10 AA386233 149 bp mRNA EST 21-APR-1997
LOCUS EST07409 Hippocampus Homo sapiens CDNA 3' end, mRNA sequence.
DEFINITION AA386233
ACCESSION AA386233
NID 92038559
VERSION AA386233.1 GI:2038559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 149)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulcher, R.A.,
 White, C.J., Lee, N.H., Kirkness, E.F., Weinstock, R.G., Gocayne, J.D.,
 Bult, C.J., Luetjens, C.M., Blake, J.A., Brandon, R.C., Mao, W.-J.,
 White, O., Sutton, R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
 Claydon, R.A., Cline, R., Fritchman, J.L., Geoghegan, N.S.,
 Fitzgerald, L.M., Fitzhugh, W.M., Hanchard, J.L., Hinkley, P.S., Jr.,
 Glodde, A., Gish, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
 Kelley, J.M., Kelley, J.C., Liu, L.-L., Marra, S.M., Merrick, J.M.,
 Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
 Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dimke, D., Feng, D.-F., Ferrel, A., Fischer, C., Hastings, G.A.,
 He, W.-W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
 Kozak, D.L., Kunsch, C., Hung, J., Xu, C., Yu, G.L., Ruben, S.M.,
 Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M.,
 Dillon, P.J., Fannon, M.R., Rosen, C.A., Hessel, W.A., Fields, C.,
 Fraser, C.M. and Venter, J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT On Sep 21, 1992 this sequence version replaced gi:2762272.
 Other ESTs: THC109664
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3016699056
 Fax: 3016699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tcdb/hgi/hgi.html>)
 Seq primer: M13-21.

FEATURES
 SOURCE Location/Qualifiers
 1..149
 /organism="Homo sapiens"
 /note="Organ: brain; Vector: pluescript SK-; Site_1:
 EcoRI; Site_2: EcoRI"
 /db_xref="taxon:9606"
 /clone_lib="Hippocampus"
 /tissue_type="hippocampus"
 BASE COUNT 39 a 41 c 34 g 32 t 3 others
 ORIGIN

Query Match 100.0%; Score 9; DB 12; Length 149;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 TTCTGAGAA 64
 |||||
 QY 25 ttctgagaa 33

RESULT 11
 LOCUS AA930548 150 bp mRNA EST 23-APR-1998
 DEFINITION vs60d02.r1 StrataGene mouse skin (#937313) Mus musculus cDNA clone
 IMAGE:1150659 5', mRNA sequence.
 ACCESSION AA930548
 NID 93080038
 VERSION AA930548.1 GI:3080038
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 150)
 REFERENCE Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2287244.

FEATURES
 SOURCE Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:623867
 Seq primer: -28m13 rev1 ET from Amersham.
 Location/Qualifiers
 1..150
 /organism="Mus musculus"

/strain="C57BL/6"
 /note="Organ: skin; Vector: pluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 011go dt. Whole skin from 11 week old C57BL/6 female mice.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
 adaptor sequence: 5' GAATTCGGCAGCG 3' -3' adaptor
 sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'."
 /db_xref="taxon:10090"
 /map="7; 19"
 /clone="IMAGE:1150659"
 /clone_lib="StrataGene mouse skin (#937313)"
 /sex="females"
 /tissue_type="whole skin"
 /dev_stage="11 weeks old"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 39 a 28 c 33 g 50 t
 ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 21 TTCTGAGAA 29
 |||||
 QY 25 ttctgagaa 33

RESULT 12
 LOCUS AA472838 151 bp mRNA EST 18-JUN-1997
 DEFINITION vds8f07.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
 clone IMAGE:804805 5' similar to TR:EI8396 EI28396 XPMC2 PROTEIN.
 ; mRNA sequence.
 ACCESSION AA472838
 NID 92201065
 VERSION AA472838.1 GI:2201065
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 151)
 REFERENCE Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:692467.

FEATURES
 SOURCE Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:481149
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 High quality sequence stop: 1.
 Location/Qualifiers
 1..151
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally from mRNA prepared
 from 800 blastocysts. Primer: SalI(dT): cDNAs were
 5'-CGGTGACCGTCGACCGTTTTTTTTTTT-3'. cDNAs were
 cloned into the NotI/SalI sites of a pSPORT vector (Life
 Technologies). Two different size selections: B1 (larger

Inserts) and B3."
 /db_xref="taxon:10090"
 /map="959H08:1"
 /clone="IMAGE:804805"
 /clone_1lb="Knowles Solter mouse blastocyst B1"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"
 /lab="host=DH10B"

BASE COUNT 39 a 34 c 46 g 32 t

ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 46 TTCTGAGAA 54
 |||||
 25 ttctgagaa 33

RESULT 13
 LOCUS T23127 153 bp mRNA EST 04-OCT-1995
 DEFINITION CS144X chinook salmon pituitary Cyclopterus lumpus cDNA clone
 CS144X 5' end similar to somatolactin, mRNA sequence.
 ACCESSION T23127
 MID 6505948
 VERSION T23127.1 GI:505948
 KEYWORDS EST.
 SOURCE Cyclopterus lumpus.
 ORGANISM Cyclopterus lumpus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
 Scorpaeniformes; Cottoidae; Cyclopteridae; Cyclopterus.
 1 (bases 1 to 153)
 Gong, Z.
 Atlantic halibut CDNAS
 Unpublished (1994)

REFERENCE
 AUTHORS Atlantic halibut CDNAS
 TITLE Unpublished (1994)
 JOURNAL

COMMENT

Contact: Zhiyuan Gong
 School of Biological Sciences
 National University of Singapore
 Lower Kent Ridge Road, Singapore 119260
 Tel: 65-8742860
 Fax: 65-7792486
 Email: obszy@leonis.nus.edu.sg
 Seq primer: SK.

FEATURES
 Source
 Location/Qualifiers
 1..153
 /organism="Cyclopterus lumpus"
 /note="Vector: Lambda Uni-ZAP XR; Site.1: EcoRI; Site.2:
 XhoI; mRNA was purified from several pituitaries of the
 chinook salmon. cDNA was directly cloned into the lambda
 ZAPRI vector using the Stratagene's Lambda Uni-ZAP XR cDNA
 cloning system."
 /db_xref="taxon:8103"
 /clone="CS144X"
 /clone_1lb="chinook salmon pituitary"

BASE COUNT 55 a 19 c 36 g 43 t

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 153;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 TTCTGAGAA 71
 |||||
 25 ttctgagaa 33

RESULT 14
 LOCUS AA786748 156 bp mRNA EST 31-JUL-1998
 DEFINITION m5906a1.r1 Aspergillus nidulans 24hr asexual developmental and
 vegetative cDNA lambda zap library Emeritella nidulans cDNA clone

m5906a1 5', mRNA sequence.
 AA786748
 92847078
 AA786748.1 GI:2847078
 EST.
 SOURCE Emeritella nidulans.
 ORGANISM Emeritella nidulans.
 Eukaryota; Fungi; Ascomycota; Euascomycetes; Plectomycetes;
 Eurotiales; Trichocomaceae; Emeritella.
 1 (bases 1 to 156)
 Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 Prade, R., and Roe, B.
 An Aspergillus nidulans EST Database
 Unpublished (1998)
 On Jan 19, 1998 this sequence version replaced gi:2150227.

REFERENCE
 AUTHORS
 TITLE An Aspergillus nidulans EST Database
 JOURNAL
 COMMENT

Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 Seq primer: SK
 High quality sequence stop: 146.

FEATURES
 Source
 Location/Qualifiers
 1..156
 /organism="Emeritella nidulans"
 /strain="FGSC A26"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"
 /db_xref="taxon:5072"
 /map="Xq28"
 /clone="m5906a1"
 /clone_1lb="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"

BASE COUNT 47 a 27 c 38 g 44 t

ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 156;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 TTCTGAGAA 42
 |||||
 33 ttctgagaa 25

RESULT 15
 LOCUS AA801391 159 bp mRNA EST 30-APR-1998
 DEFINITION EST190888 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone
 RSPA38 3' end similar to FGF-binding protein (FGF-BP), mRNA
 sequence.
 AA801391
 92864346
 AA801391.1 GI:2864346
 EST.
 VERSION AA801391
 MID 92864346
 KEYWORDS EST.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 159)
 Lee, N.H., Glodex, A., Chandra, I., Mason, T.M., Quackenbush, J.,
 Kerlavage, A.R., and Adams, M.D.
 Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat
 Gene Index
 Unpublished (1998)
 On Sep 12, 1996 this sequence version replaced gi:1407396.
 Other_ESTs: EST190889

Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

FEATURES

source

Location/Qualifiers
1..159
/organism="Rattus sp."
/note="Organ: spleen; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"
/db_xref="ATCC (Inhost):2009482"
/db_xref="taxon:10118"
/clone="RSPAA38"
/clone_lib="Normalized rat spleen, Bento Soares"
BASE COUNT 47 a 30 c 32 g 50 t
ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 9 TTCTGAGAA 17
|||||
OY 25 ttctgagaa 33

RESULT 16

LOCUS AA165736 160 bp mRNA EST 19-DEC-1996
DEFINITION nt3f01.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
IMAGE:635545.5', mRNA sequence.

ACCESSION AA165736
NID 91743952
VERSION AA165736.1 GI:1743952

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The Mashu-HMI Mouse EST Project
JOURNML Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1324745.

Contact: Marra M/Mouse EST Project
Washu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:387537
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 138.
Location/Qualifiers
1..160
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: lymph node; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACCAATCTGAAGTGGAGCGCCGCGATCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors

FEATURES

source

Location/Qualifiers
1..160
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: lymph node; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACCAATCTGAAGTGGAGCGCCGCGATCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors

BASE COUNT 45 a 39 c 23 g 53 t
ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 75 TTCTGAGAA 83
|||||
CP 33 ttctgagaa 25

RESULT 17

LOCUS AA19466 162 bp mRNA EST 30-DEC-1997
DEFINITION zhd0b06.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone
IMAGE:414515.3', mRNA sequence.

ACCESSION AA19466
NID 92732565
VERSION AA19466.1 GI:2732565

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
Unpublished (1997)

TITLE Mashu-NCI human EST Project
JOURNML Unpublished (1997)
COMMENT On Nov 6, 1997 this sequence version replaced gi:947573.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40u13 fwd, 5' from Amersham.

FEATURES

source

Location/Qualifiers
1..162
/organism="Homo sapiens"
/note="Organ: pineal gland; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTTACCAATCTGAAGTGGAGCGCCGCGATCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was size selected, ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT73
vector (Pharmacia). Library constructed by Bento Soares
and M.Fatima Bonaldo.

/db_xref="GDB:1319786"
/db_xref="taxon:9606"
/clone="IMAGE:414515"
/clone_lib="Soares_pineal_gland_N3HPG"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 57 a 33 c 30 g 42 t

ORIGIN

Query Match	100.0%;	Score 9;	DB 17;	Length 162;
Best Local Similarity	100.0%;	Pred. No. 2.23e+03;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Db	124	TTCTCAGAA	13
Cp	33	ttctcagaa	25

RESULT	18
LOCUS	AA12257 164 bp mRNA EST 31-JAN-1997
DEFINITION	mus0a11.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA clone IMAGE:551836 5' similar to gb:U0049 Mouse mRNA for mouse melanoma antigen, complete cds (MOUSE);, mRNA sequence.
ACCESSION	AA12257

ORGANISM Mus musculus

REFERENCE
1 (pages 1 to 164)
Marré, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Apr 14, 1993 this sequence version replaced gi:692716.

Contact: Maria M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.
MGI:397684

FEATURES	Location/Qualifiers
source	1. .164

```

/organism="Mus musculus"
/note="Organ: Skin; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. From M2 cells, a highly metastatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 Kb;
unt_ZAP XR Vector: -5' adaptor sequence: 5' GATTCGCACGACG
3' -3' adaptor sequence: 5' CTCGACATTTTTTTTTTTTTTT 3'"
/db_xref="taxon:10090"
/map="1"
/clone="IMAGE:651836"
/clone_lib="Stratagene mouse melanoma (#937312)"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOAR (kanamycin resistant)"

```

LOCUS	RESULT	19	AA369389	169 bp	mRNA	EST	21-APR-1997
Db	8	TTCTCAGAA	16				
Cp	33	ttctcagaa	25				

DEFINITION	EST80794 Placenta II Homo sapiens cDNA 5' end, mRNA sequence
ACCESSION	AA369389
NID	92021708
VERSION	AA369389.1 GI:2021708
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 169)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,

Bull, C. O., Leen, R., Kirkness, E. E., Weinstock, A. V., Goodbye, J. D., White, O., Sutton, G., Blake, J. R., Brandon, R. C., Man, M.-L., C. L., Clayton, R. A., Cline, T. R., Cotton, M. D., Earle-Hughes, J., Fine, L. D., Fitzgerald, L. M., Fitzhugh, W. M., Friedman, J. L., Geoghegan, N. S., Glodde, A., Gnehm, C. L., Hanna, M. C., Hedblom, E., Hinkle, S. J., Kelley, J. M., Kelley, C. C., Liu, L. I., Marmaros, S. M., Merrick, J. M., Moreno-Palancques, R. F., McDonald, L. A., Nguyen, D. T., Pelligrino, S. M., Phillips, C. A., Ryder, S. E., Scott, J. L., Saudke, D. M., Shirley, R., Small, K. V., Spilgys, T. A., Utterback, T. R., Wetman, J. F., Li, Y., Bednarek, D. P., Cao, L., Cepeda, M. A., Coleman, T. A., Collins, E. J., Dimke, D., Feng, D. F., Fertile, A., Fischer, C., Hastings, G. A., He, W. W., Hu, J. S., Greene, J. M., Gruber, J., Hudson, P., Kim, A. K., Kayak, D. L., Kunsch, C., Hungjun, J., Li, H., Melsner, P. S., Olsen, H., Raymond, L., Wei, Y. F., Wang, J., Xu, C., Yu, G. L., Ruben, S. M., Dillion, P. J., Fannon, M. R., Rosen, C. A., Haseltine, W. A., Fields, C., Fraser, C. M., and Venter, J. C.

TITLE
Initial assessment of human gene diversity and expression patterns
based on 83 million nucleotides of cDNA sequence

JOURNAL
NATURE 377 (6547 Suppl.), 3-174 (1995)

COMMENT On Sep 12, 1996 this sequence version replaced g1:1393506-045-1707. EMBL74334

COMMENT On Sep 12, 1996 this sequence version replaced g1:1393506-045-1707. EMBL74334

COMMENT On Sep 12, 1996 this sequence version replaced g1:1393506-045-1707. EMBL74334

Other_ESTS: THCI74224

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel.: 3018699056
Fax: 3018699423
Email: arkerlavet@igf.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/cdb/hg1/hg1.html>)
Seq primer: M13 Reverse.

FEATURES	Location/Qualifiers
source	1. .169

```

/organism="Homo sapiens"
/ntoe=Organ: placenta: Vector: pbluescript SK-; site_1:
EcoRI; site_2: xhoI
/db_xref="AtCC (Inhost):174008"
/db_xref="taxon:9606"
/clone_id="placenta II"
/tissue_type="placenta"
/dev_stage="fetus"
BASE COUNT      46 a      46 c      43 g      33 t      1 others
ORIGIN

```

```

Query Match      100.0%;   Score 9;   DB 12;   Length 169;
Best Local Similarity 100.0%;   Pred. No. 2.23e+03;
Matches      9;   Conservative      0;   Mismatches      0;   Indels      0;   Gaps      0;

Db      134   TTCGAGAGA   142
QY      25    ttccgagaa   33

```

LOCUS	20	173 bp	EST	02-SEP-1997
DEFINITION	vml3102.r1 Knowles Soliter mouse blastocyst B1 Mus musculus cDNA			
ACCESSION	AA574819			
NID	92349445			
VERSION	AA574819.1 GI:2369445			

KEYWORDS
EST.
house mouse.
Mus musculus.

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 173)

AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE
The Mashu-HHMI Mouse EST Project

JOURNAL
Unpublished (1996)

COMMENT
On Sep 12, 1996 this sequence version replaced gi:1402317.

FEATURES
source
Contact: Marra M/Mouse EST Project
Mashu-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG:564106
Putative full length read
vector to vector length is 174.
Location/Qualifiers
1..173
/organism="Mus musculus"
/strain="B6D2 F1/J"
/note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally from mRNA prepared
from 800 blastocysts. Primer: SalI(dT):
5'-CGTGCAGCGTCGACCGCTTTTCTTTT-3'. CDNA were
cloned into the NotI/SalI sites of a pSPORT vector (Life
Technologies). Two different size selections: B1 (larger
inserts) and B3."
/db_xref="taxon:10090"
/clone_image="991826"
/clone_id="Knowles Solter mouse blastocyst B1"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
BASE COUNT 58 a 31 c 29 g 55 t
ORIGIN
Query Match 100.0%; Score 9; DB 15; Length 173;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 TTCACGAA 36
|||||||
Cp 33 ttctcagaa 25

RESULT 21
LOCUS AA215399 173 bp mRNA EST 13-NOV-1997
DEFINITION z196h04.f1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:583575 5',
RNA sequence.
ACCESSION AA215399
VERSION AA215399.1 GI:1815144
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 173)
AUTHORS NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT
On Sep 21, 1992 this sequence version replaced gi:276423.

CONTACT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1272 Std Error: 0.00
Seq primer: -26m13 rev2 ET from Amersham
High quality sequence stop: 135.
Location/Qualifiers
1..173
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Maril (CBER). CDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTAAGTGGAGGCGCCCTCTTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="GDB:5586895"
/db_xref="taxon:9606"
/clone_image="683575"
/clone_id="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT 39 a 40 c 49 g 41 t 4 others
ORIGIN
Query Match 100.0%; Score 9; DB 9; Length 173;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 TTCACGAA 147
|||||||
Cp 33 ttctcagaa 25

RESULT 22
LOCUS AA302027 176 bp mRNA EST 18-APR-1997
DEFINITION EST15175 Aorta endothelial cells Homo sapiens CDNA 5' end, mRNA
sequence.
ACCESSION AA302027
VERSION AA302027.1 GI:1954420
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 176)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brannon, R.C., Man, M.L.,
Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dinkar, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A.,
He, W.M., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, D., Li, H., Meissner, P.S., Olsen, H.,
Raymond, D.L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,

TITLE Fraser, C.M. and Venter, J.C.
 JOURNAL Initial assessment of human gene diversity and expression patterns
 MEDLINE based upon 83 million nucleotides of cDNA sequence
 96026280 Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT On Sep 12, 1996 this sequence version replaced gl:1407137.
 Other_ESTs: THC190585
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source Location/Qualifiers

1..176
 /organism="Homo sapiens"
 /note="Organ: aorta; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"
 /db_xref="ATCC (lnhost):114253"
 /db_xref="taxon:9606"
 /clone_lib="Aorta endothelial cells"
 /cell_type="endothelial cell"
 /dev_stage="adult"
 BASE COUNT 52 a 43 c 41 g 38 t 2 others

Query Match 100.0%; Score 9; DB 11; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 TTCTGAGAA 154
 |||||
 QY 25 ttctgagaa 33

RESULT 23
 LOCUS hbc2035 178 bp mRNA EST 28-APR-1994
 DEFINITION hbc2035 Human pancreatic islet Homo sapiens cDNA clone hbc2035 5',
 mRNA sequence.
 ACCESSION T18548
 NID 9475356
 VERSION T18548.1 GI:475356
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 178)
 Bell, G.I. and Takeda, J.
 Human pancreatic islet cDNAs
 Unpublished (1995)
 JOURNAL On Sep 21, 1992 this sequence version replaced gl:276270.
 COMMENT

Contact: Bell GI or Takeda J
 HHMI
 Univ. of Chicago
 5841 S. Maryland Ave., MC1028, Chicago IL 60637
 Tel: 3127029116
 Fax: 3127020271
 Email: g-bell@uchicago.edu
 Seq primer: SK primer.
 Location/Qualifiers

FEATURES
 source 1..178
 /organism="Homo sapiens"
 /note="Vector: Lambda ZAPIT; Site_1: Eco RI; Site_2: Xho
 I; mRNA was prepared from normal adult human islets. cDNA
 was directionally synthesized from the Xho I in the vector
 to the EcoRI site. cDNA was size fractionated to remove

sequences <1000 bp in size."
 /db_xref="taxon:9606"
 /clone="hbc2035"
 /clone_lib="Human pancreatic islet"
 BASE COUNT 56 a 26 c 30 g 66 t

Query Match 100.0%; Score 9; DB 8; Length 178;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 TTCTGAGAA 107
 |||||
 QY 25 ttctgagaa 33

RESULT 24
 LOCUS AA283862 184 bp mRNA EST 13-AUG-1997
 DEFINITION z546c05.r1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:700520 5',
 similar to SW:YCD3_YEAST P25566 HYPOTHETICAL 19.3 KD PROTEIN IN
 STE50 5' REGION.;, mRNA sequence.

ACCESSION AA283862
 NID 91928071
 VERSION AA283862.1 GI:1928071
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 184)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL On Apr 14, 1993 this sequence version replaced gl:716611.
 COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (<http://image.llnl.gov>) for further information.
 trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES
 source 1..184

/organism="Homo sapiens"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+ IgD-),
 provided by Dr. Louis M. Staudt (NCI). Dr. David Altman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 15'-TGTTACCAATCTGAAGTGGAGCGCCGCTATTTTCTTTT-
 3'. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="GDB:5739948"
 /db_xref="taxon:9606"
 /map="3"
 /clone="IMAGE:700520"
 /clone_lib="NCI-CGAP GCBI"
 /issue_type="germinal center B cell"
 /lab_host="DH10B"

BASE COUNT 52 a 43 c 48 g 41 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 10; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 136 TTCGAGAA 144
 1111111111
 Oy 25 ttctgagaa 33

RESULT 25
 LOCUS AA547150 185 bp mRNA EST 05-AUG-1997
 DEFINITION v69e10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
 IMAGE:959946.5' similar to gb:M88163 POSSIBLE GLOBAL TRANSCRIPTION
 ACTIVATOR SNF2L (HUMAN);, mRNA sequence.
 ACCESSION AA547150
 NID 92308441
 VERSION AA547150.1 GI:2308441
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 185)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The Washu-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 Washu-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:548738
 Trace considered overall poor quality
 High quality sequence stop: 1.
 Location/Qualifiers
 1..185
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /note="Organ: embryo; Vector: Bluescribe (modified);
 Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
 mRNA prepared from 13,500 2-cell stage embryos. Primer:
 SalI(dT): 5'-CGGTGACCGACGACCGTGTGTGTGT-3'. cDNAs
 were cloned into the MluI/SalI sites of a modified
 Bluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 kb."
 /db_xref="taxon:10090"
 /clone_lib="Knowles Solter mouse 2 cell"
 /clone="IMAGE:959946"
 /tissue_type="embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"

BASE COUNT 74 a 28 c 48 g 35 t

ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 185;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 60 TTCGAGAA 68
 1111111111
 Cp 33 ttctgagaa 25

RESULT 26
 LOCUS A1046096 187 bp mRNA EST 07-JUL-1998
 DEFINITION TENU1182 T. cruzi epimastigote normalized cDNA library Trypanosoma

ACCESSION A1046096
 NID g3293994
 VERSION A1046096.1 GI:3293994
 KEYWORDS EST.
 SOURCE Trypanosoma cruzi.
 ORGANISM Trypanosoma cruzi.
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma; Schizotrypanum.
 REFERENCE 1 (bases 1 to 187)
 AUTHORS Porcel,B., Tran,A-N., Andersson,B., Nyarady,Z., Urmenyi,T.P.,
 Rondinelli,E., Pettersson,U. and Aslund,L.
 TITLE Generation of Expressed Sequence Tags from Trypanosoma cruzi
 JOURNAL Unpublished (1998)
 COMMENT On Sep 1, 1995 this sequence version replaced.

FEATURES
 source
 1..187
 /organism="Trypanosoma cruzi"
 /strain="Cl-Brenner"
 /note="cDNA library constructed with oligo dt primed
 epimastigote mRNA and cloned in pTc118D phagemid with
 modified polylinker (PHARMACIA)"
 /db_xref="taxon:5693"
 /map="5"
 /clone="20f19"
 /clone_lib="T. cruzi epimastigote normalized cDNA library"
 /cell_type="epimastigote"

BASE COUNT 33 a 59 c 40 g 55 t

ORIGIN

Query Match 100.0%; Score 9; DB 22; Length 187;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 29 TTCGAGAA 37
 1111111111
 Cp 33 ttctgagaa 25

RESULT 27
 LOCUS AA488722 187 bp mRNA EST 05-MAR-1998
 DEFINITION aa95e01.t1 Stragagene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:83064.5' similar to contains Ll.t1 Ll repetitive element ;,
 mRNA sequence.
 ACCESSION AA488722
 NID g2218324
 VERSION AA488722.1 GI:2218324
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 187)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
 Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
 Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
 TITLE Washu-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1392955.
 Contact: Wilson RK
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1736 Std Error: 0.00
Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 144.

FEATURES

source

1. 187

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

/db_xref="taxon:9606"

/clone="IMAGE:839064"

/clone_1lb="Stratagene fetal retina 937202"

/sex="mixed"

/lab_host="SOLR (kanamycin resistant)"

/lab_host="SOLR (kanamycin resistant)"

BASE COUNT

73 a 34 c 29 g 51 t

ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 TTCTCAGAA 40
|||||
Cp 33 ttctcagaa 25

RESULT

LOCUS AA911029 197 bp mRNA EST 13-APR-1998
DEFINITION ORF1f05.s1 NCI CGAP GC4 Homo sapiens cDNA clone IMAGE:1518465 3'
similar to TR:016342 016342 PDCD2-PROGRAMMED CELL DEATH-2/RP8

HOMOLOG. ; mRNA sequence.

AA911029
93050319
AA911029.1 GI:3050319

KEYWORDS
EST.
human.

SOURCE

human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 197)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL On Jan 19, 1998 this sequence version replaced gi:2284616.

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bdrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 167.

FEATURES
source

1. 197

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled

BASE COUNT

43 a 54 c 49 g 51 t

ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 88 TTCTCAGAA 96
|||||

Oy 25 ttctcagaa 33

RESULT

LOCUS AA301402 197 bp mRNA EST 18-APR-1997
DEFINITION EST14416 Testis tumor Homo sapiens cDNA 5' end, mRNA sequence.

AA301402
91953809
AA301402.1 GI:1953809

KEYWORDS
EST.
human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 197)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fullmer, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Meisnick, R.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitch, W.M., Fitchman, J.L., Geoghagen, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palacios, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bedard, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hung, J., Li, H., Meisner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Pannon, M.R., Rosen, C.A., Hasettline, W.A., Fields, C.,
Fraser, C.M., and Venter, J.C.

AUTHORS

JOURNAL

Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1405016.

TITLE

JOURNAL

Medline 96026280

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423

FEATURES

source

1. 197

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled

germ cell tumors, and was then primed with a Not I -
Oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo.

Seq primer: M3 Reverse

Location/Qualifiers

1. 197

```

/note="Organ: testis; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="ATCC (Inhost):191527"
/db_xref="taxon:9606"
/mab="8"
/clone_lib="Testis tumor"
/dev_stage="adult"
BASE COUNT      43 a      45 c      35 g      73 t      1 others
ORIGIN
Query Match      100.0%; Score 9; DB 11; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      5 TTCTCAGAA 13
|||||
Cp      33 ttctcagaa 25

RESULT      30
LOCUS      AA806861      198 bp      mRNA      EST      25-MAR-1998
DEFINITION      OC29e10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1351146 3',
                mRNA sequence.
ACCESSION      AA806861
VERSION      92876437
KEYWORDS      AA806861.1 GI:2876437
SOURCE      EST.
ORGANISM      human.
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 198)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On Jan 19, 1998 this sequence version replaced gi:2150271.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLN at:
www.bio.livl.gov/dbtrp/Image/Image.html

Insert Length: 829 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 183.
Location/Qualifiers
1. 198
/organism="Homo sapiens"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germlinal center B cells by flow sorting (CD20+ IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCAATTTTCTTTTCTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1351146"

```

```

/clone_lib="NCI CGAP GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT      62 a      42 c      43 g      51 t
ORIGIN
Query Match      100.0%; Score 9; DB 18; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      72 TTCTCAGAA 80
|||||
Cp      33 ttctcagaa 25

RESULT      31
LOCUS      AA381661      199 bp      mRNA      EST      21-APR-1997
DEFINITION      EST95014 Activated T-cells I Homo sapiens cDNA 5' end similar to
                similar to protein kinase PKM2, mRNA sequence.
ACCESSION      AA381661
VERSION      92034206
KEYWORDS      AA381661.1 GI:2034206
SOURCE      EST.
ORGANISM      human.
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 199)
AUTHORS      Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
                Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
                White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Val,C.,
                Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
                Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Georgagen,N.S.,
                Gloder,A., Gnehm,C.L., Hanna,M.C., Heddlom,E., Hinkle,P.S.Jr.,
                Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
                Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
                Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
                Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
                Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
                Dinke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
                He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
                Korak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
                Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
                Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
                Fraser,C.M. and Venter,J.C.
                Initial assessment of human gene diversity and expression patterns
                based upon 83 million nucleotides of cDNA sequence
                Nature 377 (6547 suppl), 3-174 (1995)
JOURNAL      MEDLINE
COMMENT      On Sep 12, 1996 this sequence version replaced gi:1406865.
                Other ESTs: THC100111
                Contact: Kerlavage, AR
                Bioinformatics
                The Institute for Genomic Research
                9712 Medical Center Drive, Rockville, MD 20850 USA
                Tel: 3016699056
                Fax: 3016699423
                Email: atkerlavet@igf.org
                For clone availability, additional sequence and expression
                information related to this EST, please check the TIGR Human Gene
                Index (http://www.tigr.org/tdb/hgi/hgi.html)
                Seq primer: M13 Reverse.
                Location/Qualifiers
                1. 199
                /organism="Homo sapiens"
                /note="Vector: pBluescript SK-; Site_1: EcorI; Site_2:
                XhoI"
                /db_xref="ATCC (Inhost):186260"
                /db_xref="taxon:9606"
                /clone_lib="Activated T-cells I"
                /cell_type="T-lymphocyte"
                /dev_stage="adult"
BASE COUNT      63 a      40 c      46 g      49 t      1 others

```

ORIGIN

Query Match 100.0% Score 9; DB 12; Length 199;
 Best Local Similarity 100.0%;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 180 TTCTCAGAA 188
 |||||
 25 ttctcagaa 33

RESULT 32
 LOCUS AA367158 199 bp mRNA EST 21-APR-1997
 DEFINITION EST978389 Pancreas tumor III Homo sapiens CDNA 5' end, mRNA
 sequence.
 ACCESSION AA367158
 NID 92019644
 VERSION AA367158.1 GI:2019644
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 199)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Mali,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.T., Geoghagen,N.S.,
 Glodok,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
 Dinko,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
 He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meltsner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M., and Venter,J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 suppl), 3-174 (1995)

JOURNAL MEDLINE
 96026280
 On Sep 12, 1996 this sequence version replaced gi:1290202.
 Other-ESTs: THCI7896
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavet@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. 199
 /organism="Homo sapiens"
 /note="Organ: pancreas; Vector: pBluescript SK-; Site:1;
 EcoRI: Site:2; XhoI"
 /db_xref="ATCC (Inhost):172019"
 /db_xref="taxon:9606"
 /clone_id="Pancreas tumor III"
 /dev_stage="adult"

BASE COUNT 61 a 33 c 24 g 81 t

ORIGIN

Query Match 100.0% Score 9; DB 12; Length 199;
 Best Local Similarity 100.0%;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 187 TTCTCAGAA 195
 |||||
 33 ttctcagaa 25

RESULT 33
 LOCUS 219882 201 bp mRNA EST 07-FEB-1995
 DEFINITION HSAAAMWK H, Human adult Brain Cortex tissue Homo sapiens cDNA,
 mRNA sequence.
 ACCESSION 219882
 NID 926627
 VERSION 219882.1 GI:26627
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 201)
 MRC Human Genome Mapping Project Resource Centre.
 The UK-HGMP cDNA program
 Unpublished (1993)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

CONTACT: MRC Human Genome Mapping Project Resource Centre
 Clinical Research Centre
 Watford Road, Harrow, Middlesex HA1 3UJ, U.K.
 Email: biocentre@hgm.mrc.ac.uk
 Single read.
 Location/Qualifiers
 1. 201
 /organism="Homo sapiens"
 /note="Vector: gtl1; clone library=H, Human adult Brain
 Cortex tissue; Cloning vector is gtl1."
 /db_xref="taxon:9606"
 /clone_id="H, Human adult Brain Cortex tissue"

BASE COUNT 61 a 53 c 42 g 44 t

ORIGIN

Query Match 100.0% Score 9; DB 8; Length 201;
 Best Local Similarity 100.0%;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 102 TTCTCAGAA 110
 |||||
 33 ttctcagaa 25

RESULT 34
 LOCUS AA591051 203 bp mRNA EST 16-SEP-1997
 DEFINITION V147C10.s1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
 IMAGE:975378 3' similar to SW:R111_HUMAN P39026 605 RBOSOMAL
 PROTEIN L11. ", mRNA sequence.
 ACCESSION AA591051
 NID 92404364
 VERSION AA591051.1 GI:2404364
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 203)
 Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geiselberg,K., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 On Sep 12, 1996 this sequence version replaced gi:1394211.

JOURNAL
 COMMENT

CONTACT: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine


```

DEFINITION   hbc032B Human pancreatic islet Homo sapiens cDNA clone hbc032B
              3' end, mRNA sequence.
ACCESSION    T10615
NID          9390769
VERSION      T10615.1 GI:390769
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE    1 (bases 1 to 205)
AUTHORS      Takeda,J., Yano,H., Eng,S., Zeng,Y. and Bell,G.I.
TITLE        A molecular inventory of human pancreatic islets: sequence analysis
              of 1000 cDNA clones
JOURNAL      Hum. Mol. Genet. 2, 1793-1798 (1993)
MEDLINE      94108427
COMMENT      Contact: Bell GI or Takeda J
              HHMI
              Univ. of Chicago
              5841 S. Maryland Ave., MC1028, Chicago IL 60637
              Tel: 3127029116
              Fax: 3127020271
              Email: g-bell@uchicago.edu
              Seq primer: T7 primer.
FEATURES
  source
    1..205
    Location/Qualifiers
      /organism="Homo sapiens"
      /note="Vector: Lambda ZAPIT; Site_1: Eco RI; Site_2: Xho
      I; mRNA was prepared from normal adult human islets. cDNA
      was directionally synthesized from the Xho I in the vector
      to the EcoRI site. cDNA was size fractionated to remove
      sequences <1000 bp in size."
      /db_xref="taxon:9606"
      /clone="hbc032B"
      /clone_1lb="Human pancreatic islet"
BASE COUNT   79 a      34 c      28 g      64 t
ORIGIN
Query Match      100.0%; Score 9; DB 8; Length 205;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db              56 TTCTGAGA 64
              |||||||
Oy              25 ttctgagaa 33

RESULT 38
LOCUS      AA546389      207 bp      mRNA      EST      05-AUG-1997
DEFINITION v64502.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:955403 5' similar to gb:D10049 Mouse mRNA for mouse melanoma
antigen, complete cds (MOUSE);, mRNA sequence.
ACCESSION  AA546389
NID         92307680
VERSION     AA546389.1 GI:2307680
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 207)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE       The WashU-HMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     On Sep 21, 1992 this sequence version replaced gi:279367.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project

```

```

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:548195
Putative full length read
vector to vector length is 208.
FEATURES
  source
    1..207
    Location/Qualifiers
      /organism="Mus musculus"
      /note="Organ: blood; Vector: pBluescript SK-; Site_1:
      EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
      01190 dt. M30 CD4+ cells. Average insert size: 1.0 Kb;
      Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGACGAC
      3' -3' adaptor sequence: 5' CTCGACGTTTTTTTTTTTTTTT 3'."
      /db_xref="taxon:10090"
      /map="6"
      /clone="IMAGE:959403"
      /clone_1lb="Stratagene mouse Tcell 937311"
      /tissue_type="Tcell"
      /dev_stage="M30 CD4+ cells"
      /lab_host="SOLR (kanamycin resistant)"
BASE COUNT   75 a      49 c      54 g      29 t
ORIGIN
Query Match      100.0%; Score 9; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db              9 TTCTCAGA 17
              |||||||
CP              33 ttctcagaa 25

RESULT 39
LOCUS      AA146148      208 bp      mRNA      EST      11-FEB-1997
DEFINITION m886f02.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
clone IMAGE:555819 5' similar to gb:D10049 Mouse mRNA for mouse
melanoma antigen, complete cds (MOUSE);, mRNA sequence.
ACCESSION  AA146148
NID         91715520
VERSION     AA146148.1 GI:1715520
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 208)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE       The WashU-HMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1395424.
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:360467
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
FEATURES

```



```

source
1. 208
/oranism="Mus musculus"
/note="Organ: Skin; Vector: pBluescript SK-; Site:1;
ECORI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. From M2 cells, a highly metastatic derivative of
the K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTCCTTTTTCCTTTT 3'"
/db_xref="taxon:10090"
/clone_lib="IMAGE:585819"
/clone_lib="Stratagene mouse melanoma (#937312)"
/tissue_type="melanoma"
/dev_stage="M2 cells"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT      70 a      53 c      52 g      33 t
ORIGIN

Query Match      100.0%; Score 9; DB 9; Length 208;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      8 TTCTCAGAA 16
|||||
Cp      33 ttctcagaa 25

RESULT 40
LOCUS      217407      209 bp      mRNA      EST      09-NOV-1992
DEFINITION HSHHE0117 Stratagene cDNA library Human heart, cat#936208 Homo
sapiens cDNA clone HE0117, mRNA sequence.
ACCESSION      217407
NID      930553
VERSION      217407.1 GI:30553
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 209)
AUTHORS      Aufiray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
Devignes,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B.,
Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
Sebastiani-Kabakchis,C. and Tessier,A.
TITLE      IMAGE: molecular integration of the analysis of the human genome
and its expression
JOURNAL      C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE      95377534
COMMENT      Contact: Genzentrum Muenchen
Laboratorium fuer molekulare Biologie
Am Klopferstutz 18a, 8033 Martinsried, Germany
Email: obermaier@vms.blochem.mpg.de
single read.

FEATURES
Source
1. 209
Location/Qualifiers
/oranism="Homo sapiens"
/note="Vector: pBluescript SK(+); Human heart cDNA
library. Cloning vector pBluescript SK(+)"
/db_xref="GDB:0057230E"
/db_xref="taxon:9606"
/clone_lib="HE0117"
/clone_lib="Stratagene cDNA library Human heart,
cat#936208"

BASE COUNT      29 a      48 c      49 g      83 t
ORIGIN

Query Match      100.0%; Score 9; DB 8; Length 209;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      138 TTCTCAGAA 146
|||||
0y      25 ttctcagaa 33

```

```

RESULT 41
LOCUS      AA796019      210 bp      mRNA      EST      09-FEB-1998
DEFINITION V393C12.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:1148662 5' similar to gp:DI0049 Mouse mRNA for mouse melanoma
antigen, complete cds (MOUSE);, mRNA sequence.
ACCESSION      AA796019
NID      92858974
VERSION      AA796019.1 GI:2858974
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 210)
AUTHORS      Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE      The WashU-HMNI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      On Jan 17, 1998 this sequence version replaced gi:2044132.

Contact: Marra M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:621870
Putative full length read
vector to vector length is 211
Seq primer: -28m13 rev1 ET from Amersham.

FEATURES
Source
1. 210
Location/Qualifiers
/oranism="Mus musculus"
/note="Organ: blood; Vector: pBluescript SK-; Site:1;
ECORI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTCCTTTTTCCTTTT 3'"
/db_xref="taxon:10090"
/map="8"
/clone_lib="IMAGE:1148662"
/clone_lib="Stratagene mouse Tcell 937311"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT      75 a      50 c      56 g      29 t
ORIGIN

Query Match      100.0%; Score 9; DB 18; Length 210;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      9 TTCTCAGAA 17
|||||
Cp      33 ttctcagaa 25

RESULT 42
LOCUS      AA357912      210 bp      mRNA      EST      21-APR-1997
DEFINITION EST66720 Fetal Lung III Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION      AA357912
NID      92010232
VERSION      AA357912.1 GI:2010232
KEYWORDS      EST.
SOURCE      human.

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 210)
AUTHORS Adams M.D., Soares M.B., Kerlavage A.R., Fields C. and Venter J.C.
TITLE Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library
JOURNAL Nature Genet. 4, 373-380 (1993)
MEDLINE 94004965
COMMENT On Sep 12, 1996 this sequence version replaced g1:1397416.
Other ESTs: THC101658
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability: additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1..210
/organism="Homo sapiens"
/note="Organ: Lung; Vector: pBluescript SK-; Site_1: ECORI; Site_2: XhoI."
/db_xref="ATCC (lnhost):162236"
/db_xref="taxon:9606"
/clone_lib="Fetal Lung III"
/dev_stage="fetus"
BASE COUNT 82 a 31 c 45 g 51 t 1 others
ORIGIN
Query Match 100.0%; Score 9; DB 12; Length 210;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 TTCACGAA 79
|||||||
Cp 33 ttctcagaa 25

RESULT 43
LOCUS AA824791 210 bp mRNA EST 19-FEB-1998
DEFINITION CT093.UNIT Tomato Leaf cDNA from cv. VFNT cherry Lycopersicon esculentum cDNA clone CT093, mRNA sequence.
ACCESSION AA824791
NID 92897269
VERSION AA824791.1 GI:2897269
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; Asteridae; Solanaceae; Solanales; Solanaceae; Solanum; Potaceae; Lycopersicon.
1 (bases 1 to 210)
Ganal,M.W., Czllnal,R., Hannappel,U., Kloos,D.-U., Polley,A. and Ling,H.-Q.
TITLE Sequencing of cDNA clones from the genetic map of tomato (Lycopersicon esculentum)
JOURNAL Unpublished (1998)
COMMENT On Apr 14, 1993 this sequence version replaced g1:692655.
Contact: Ganal, M.W.
Dept. of CytoGenetics
Institute of Plant Genetics and Crop Plant Research
Corrensstr. 3, D-06466 Gatersleben, Germany
Tel: 49-39482-5346
Fax: 49-39482-5137
Email: ganal@emendel.ipk-gatersleben.de
Seq primer: M13-20.

FEATURES
source Location/Qualifiers
1..210
/organism="Lycopersicon esculentum"
/cultivar="VFNT cherry"
/note="Organ: Leaf; Vector: pBluescript"
/db_xref="taxon:4081"
/clone_lib="Tomato Leaf cDNA from cv. VFNT cherry"
/tissue_type="epidermis"
BASE COUNT 38 a 35 c 48 g 89 t
ORIGIN
Query Match 100.0%; Score 9; DB 18; Length 210;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 TTCACGAA 24
|||||||
Cp 33 ttctcagaa 25

RESULT 44
LOCUS AA653799 212 bp mRNA EST 23-DEC-1997
DEFINITION nt02009.g1 NCI_CGAP_Lym3 Homo sapiens cDNA clone IMAGE:1191929 3' similar to TR:G189397 G189397 HYPOTHETICAL 33.4 KD PROTEIN. ; mRNA sequence.
ACCESSION AA653799
NID 92589953
VERSION AA653799.1 GI:2589953
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 212)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced g1:1393627.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.
Tissue Procurement: Mark Raffeld, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Kitzman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrp/image/image.html
Insert Length: 648 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers
1..212
/organism="Homo sapiens"
/note="Organ: lymph node; Vector: Bluescript SK-; Site_1: ECORI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pooled lymphomas. 5' adaptor sequence: 5' CCAATCGGCGACGAG 3' 3' adaptor sequence: 5' CCGAGATTTTTTTTTTTTTT 3' Average insert size: 0.9 kb."
/db_xref="taxon:9606"
/map="7q36.1:18pE"
/clone_image:1191929"
/clone_lib="NCI_CGAP_Lym3"

/tissue-type="lymphoma"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 53 a 61 c 60 g 38 t

Query Match 100.0%; Score 9; DB 16; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 140 TTCTCAGAA 148
|||||
ttctcagaa 33

RESULT 45
LOCUS AA218212 213 bp mRNA EST 06-FEB-1997
DEFINITION mv58g03.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone
IMAGE:659237 5' similar to gb:M62419 Mouse clathrin-associated
protein (MOUSE), mRNA sequence.

ACCESSION AA218212
NID 91827204
VERSION AA218212.1 GI:1827204
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 213)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393876.

FEATURES
SOURCE

1. 213
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: whole fetus; Vector: p7T73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGAGCGCGCTATTTTATTTTATTTT
3'] on total mouse RNA [provided by Minoru Ko, Wayne
State Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="IMAGE:659237"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue-type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"

BASE COUNT 38 a 63 c 53 g 59 t

Query Match 100.0%; Score 9; DB 9; Length 213;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 92 TTCTCAGAA 100
|||||
ttctcagaa 25

CP 33 ttctcagaa 25

RESULT 46
LOCUS AA251978 214 bp mRNA EST 13-AUG-1997
DEFINITION zs11f10.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684907 5',
mRNA sequence.

ACCESSION AA251978
NID 91886994
VERSION AA251978.1 GI:1886994
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 214)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index.
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:692655.

FEATURES
SOURCE

1. 214
Location/Qualifiers
/organism="Homo sapiens"
/note="Vector: p7T73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGAGCGCGCTATTTTATTTTATTTT
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:684907"
/clone_lib="NCI_CGAP_GCB1"
/tissue-type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 64 a 33 c 39 g 78 t

Query Match 100.0%; Score 9; DB 10; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 133 TTCTCAGAA 141
|||||
ttctcagaa 25

CP 33 ttctcagaa 25

RESULT 47
LOCUS AA444159 217 bp mRNA EST 03-JUN-1997
DEFINITION zvs1h05.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:757209
5', mRNA sequence.

ACCESSION AA444159

NID 92156834
 VERSION AA444159.1 GI:2156834
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 217)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenger, K., Stepien, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 Washu-Merck EST Project 1997
 Unpublished (1997)
 On Apr 14, 1993 this sequence version replaced gi:593545.

TITLE
 JOURNAL
 COMMENT

CONTACT: Wilson R
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Putative full length read
 The vector to vector length is 218
 Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
 source
 Location/Qualifiers
 1..217
 /organism="Homo sapiens"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer (5'
 TGTACCAATCTGAGTGGAGCGGCCCAATTTTCTTTTCTTTT 3')
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cots, and was
 constructed by Bento Soares and M. Fatima Bonalao."
 /db_xref="GDB:597782"
 /db_xref="taxon:9606"
 /clone="IMAGE:757209"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /lab_host="DH10B"

BASE COUNT 52 a 49 c 50 g 66 t
 ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 217;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 143 TTCTGAGAA 151
 |||||
 OY 25 ttctgagaa 33

RESULT 48
 LOCUS F08804 220 bp mRNA EST 23-FEB-1995
 DEFINITION HSC2TA072 normalized infant brain cDNA Homo sapiens cDNA clone
 c-2ta07 3', mRNA sequence.
 ACCESSION F08804
 NID F08804
 VERSION F08804.1 GI:677960
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 220)
 Auffray, C., Behar, G., Bols, F., Bouchier, C., da Silva, C.,

TITLE
 JOURNAL
 MEDLINE
 COMMENT

CONTACT: Genethon
 GenexPress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'interationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read, removed at sequence 5' end
 GenexPress_library_id: C; GenexPress_sequence_id: aic-2ta07
 Insert Length: 1056 Std Error: 0.00
 Seq primer: (-21)m13 universal
 High quality sequence stop: 362.
 Location/Qualifiers
 1..220
 /organism="Homo sapiens"
 /note="Organ: brain; Vector: lafmid BA; Site 1: HindIII;
 Site 2: NotI; sex=Female; dev_stage=3 months old;
 isolate=muscular atrophy patient; tissue_type=total
 brain; total mRNA was oligo(dT) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B.Souares, Psychiatry
 Dept. Columbia University, USA. Normalization method:
 Bento Soares, P.N.A.S in press"
 /db_xref="taxon:9606"
 /map="18"
 /clone_lib="normalized infant brain cDNA"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"

BASE COUNT 82 a 33 c 42 g 61 t 2 others
 ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 148 TTCTGAGAA 156
 |||||
 Cp 33 ttctgagaa 25

RESULT 49
 LOCUS AA883770 221 bp mRNA EST 04-JAN-1999
 DEFINITION A159a06 s1 Soares_NFL.T.GBC.S1 Homo sapiens cDNA clone
 IMAGE:1461586 3', mRNA sequence.
 ACCESSION AA883770
 NID A2993300
 VERSION AA883770.1 GI:2993300
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 221)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Apr 14, 1993 this sequence version replaced gi:593622.

TITLE
 JOURNAL
 COMMENT

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 960 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 182.
location/Qualifiers

FEATURES

source

1. 221
/organism="Homo sapiens"
/note="Organ: pooled; Vector: pTR73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-GCAP GC81) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/map_19: 806E08; 13: 13q22.3-13q31.3; 13p11.2-13q11"
/clone_11b="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"

BASE COUNT

88 a 31 c 46 g 56 t

ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 TTCGAGAA 43
|||||
OY 25 ttctgagaa 33

RESULT

50 AA881592 222 bp mRNA EST 26-MAR-1998

LOCUS

DEFINITION vx20611.t1 Soares 2NBMT Mus musculus cDNA clone IMAGE:1265036 5',
AA881592
Accession

ACCESSION

AA881592

NID

92990902

VERSION

AA881592.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 222)

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.Warr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

The Washu-HMI Mouse EST Project

Unpublished (1996)

On Sep 12, 1996 this sequence version replaced gi:1407185.

TITLE

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project

Washu-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:667588

Seq primer: -28m13 rev2 ET from Amersham.

Location/Qualifiers

1. 222

/organism="Mus musculus"

/strain="C57BL/6J"

BASE COUNT

73 a 50 c 59 g 40 t

ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 65 TTCGAGAA 73
|||||
OY 25 ttctgagaa 33

RESULT

51 T12418 222 bp mRNA EST 28-NOV-1994

LOCUS

DEFINITION A275F Heart Homo sapiens cDNA clone A275, mRNA sequence.
T12418
Accession

ACCESSION

T12418

NID

9597105

VERSION

T12418.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 222)

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.Liew, C.C., Hwang, D.M., Fung, Y.W., Laurensen, C., Cukerman, E.,
Teul, S., and Lee, C.Y.A catalogue of genes in the cardiovascular system as identified by
expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 91, 10645-10649 (1994)

95024171

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Liew CC

Department of Laboratory Medicine and Pathobiology

University of Toronto

Banting Institute, 100 College St., Toronto, Ontario, M5G1L5

Tel: 4169788758

Fax: 4169785650

Email: liewcc@utcc.utoronto.ca

Seq primer: GCGGCGAGACGCTCGAGACC.

Location/Qualifiers

1. 222

/organism="Homo sapiens"

/note="Vector: Lambda gtl1; Site_1: EcoRI; Site_2: EcoRI"

/db_xref="taxon:9606"

/clone_11b="Heart"

/lab_host="E.coli Y1090"

BASE COUNT

57 a 55 c 65 g 45 t

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 TTCGAGAA 47
|||||

CP 33 ttctcagaa 25

RESULT 52 AA904956 223 bp mRNA EST 08-APR-1998
LOCUS 0181c09.s1 Soares_NFL.T.GBC.S1 Homo sapiens CDNA clone
DEFINITION IMAGE:1504720 3', mRNA sequence.

ACCESSION AA904956
NID 93040079
VERSION AA904956.1 GI:3040079
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 223)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:692972.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 69.
Location/Qualifiers
1..223
/organism="Homo sapiens"
/note="Organ: pooled. Vector: pYT13D-Pac (Pharmacia) with
a modified polylinker. Site_1: Not I; Site_2: EcoRI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaudo."
/db_xref="taxon:9606"
/map="MMU15C3-C4 region; MMU16C3-C4 region"
/clone_lib="Soares_NFL.T.GBC.S1"
/clone_1lb="1504720"
/lab_host="DH10B"
/lab_host="DH10B"

BASE COUNT 94 a 26 c 45 g 58 t

Query Match 100.0%; Score 9; DB 20; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 70 TTCTGAGAA 78
|||||||
QY 25 ttctcagaa 33

RESULT 53 AA603689 223 bp mRNA EST 08-OCT-1997
LOCUS np22h10.s1 NCI-CGAP.GC5 Homo sapiens CDNA clone IMAGE:1117123 3'
DEFINITION Similar to SW:SI0C-RABIT P24480 CALGIZANIN; mRNA sequence.
ACCESSION AA603689
NID 92437550
VERSION AA603689.1 GI:2437550
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 223)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397971.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/db/rp/image/image.html

Insert length: 584 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 218.
Location/Qualifiers
1..223
/organism="Homo sapiens"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI. Cloned unidirectionally. Primer: Oligo dT. Mixed
germ cell tumors. 5' adaptor sequence: 5' GATTTCGCACGAG
3' 3' adaptor sequence: 5' CTCGAGTGTGTTTTTTTTTTTTTTT 3'
Average insert size: 0.7 kb."
/db_xref="taxon:9606"
/map="21"
/clone="IMAGE:1117123"
/clone_1lb="NCI-CGAP.GC5"
/tissue_type="germ cell tumor"
/lab_host="SOUL (kanamycin resistant)"

BASE COUNT 56 a 49 c 52 g 66 t

Query Match 100.0%; Score 9; DB 15; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 TTCTGAGAA 22
|||||||
QY 25 ttctcagaa 33

RESULT 54 AA646725 223 bp mRNA EST 28-OCT-1997
LOCUS ve46h02.r1 Soares mouse mammary gland NDMG Mus musculus CDNA clone
DEFINITION IMAGE:821235 5' similar to WP:F44E2.6 CE00184 TRANSCRIPTION
REGULATION; mRNA sequence.
ACCESSION AA646725
NID 92573154
VERSION AA646725.1 GI:2573154
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 223)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucab, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jun 18, 1996 this sequence version replaced gi:1366589.

CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:489515

Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES

Source

1..223
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTACCAATCTGAGTGGAGCGCCGCGAATGCTTTTCTTTTCTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
/db_xref="taxon:10090"
/clone="IMAGE:821039"
/clone_lib="Soares mouse mammary gland NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 62 a 50 c 60 g 51 t
ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 166 TTCTGAGAA 174

Qy 25 ttctgagaa 33

RESULT 55
LOCUS AA368477 224 bp mRNA EST 21-APR-1997
DEFINITION EST19980 Placenta I Homo sapiens cDNA similar to
NADH-ubiquinone oxidoreductase, B22 subunit, mRNA sequence.
ACCESSION AA368477
NID 92020921
VERSION AA368477.1 GI:2020921
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 224)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weissstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'wa,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghegan,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,
Small,R.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bedarick,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dincke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,

TITLE

JOURNAL
MEDLINE
COMMENT
On Sep 12, 1996 this sequence version replaced gi:1392814.
Other ESTs: TNC128658
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES

source

1..224
Location/Qualifiers
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: EcoRI"
/db_xref="ATCC (Inhost):173295"
/db_xref="ATCC (Inhost):9606"
/clone_lib="Placenta I"
/tissue_type="Placenta"
/dev_stage="fetuss"
BASE COUNT 64 a 53 c 58 g 48 t 1 others
ORIGIN

Query Match 100.0%; Score 9; DB 12; Length 224;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 166 TTCTGAGAA 176

Qy 25 ttctgagaa 33

RESULT 56
LOCUS AA301938 226 bp mRNA EST 18-APR-1997
DEFINITION EST15028 Aorta endothelial cells Homo sapiens cDNA 5' end, mRNA
sequence.
ACCESSION AA301938
NID 91954301
VERSION AA301938.1 GI:1954301
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 226)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weissstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'wa,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghegan,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,
Small,R.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bedarick,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dincke,D., Feng,D.-F., Ferrite,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruten,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
Initial assessment of human gene diversity and expression patterns

TITLE

JOURNAL
MEDLINE
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
On Sep 12, 1996 this sequence version replaced gi:1407197.
COMMENT
Other ESTs: THC165580
Contact: Kerlavage, AR
Bioinformatics for Genomic Research
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igf.org

For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgl/hgl.html>)
Seq primer: M13 Reverse.

FEATURES
SOURCE
Location/Qualifiers
1..226

/organism="Homo sapiens"
/note="Organ: aorta; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="ATCC (inhost):114131"
/db_xref="taxon:9606"
/clone_lib="Aorta endothelial cells"
/cell_type="endothelial cell"
/dev_stage="adult"
BASE COUNT 62 a 41 c 75 g 44 t 4 others
ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 226;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 104 TTCTCGAA 112
|||||||
Cp 33 ttctcgaa 25

RESULT 57
LOCUS AA945908 228 bp mRNA EST 16-JUN-1998
DEFINITION EST201407 Normalized rat lung, Bento Soares Rattus sp. cDNA clone
RLUAM80 3' end, mRNA sequence.
ACCESSION AA945908
NID G3105824
VERSION AA945908.1 GI:3105824
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 228)
Lee, N.H., Glodet, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat
Gene Index
Unpublished (1998)
On Jan 17, 1998 this sequence version replaced gi:2044117.
COMMENT

CONTACT: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers
1..228

FEATURES
SOURCE
/organism="Rattus sp."
/note="Organ: lung; Vector: pRT73pac; Site_1: EcorI;
Site_2: NotI"
/db_xref="ATCC (inhost):2015180"
/db_xref="taxon:10118"

BASE COUNT 68 a 47 c 46 g 67 t
ORIGIN
/map="8"
/clone="RLUAM80"
/clone_lib="Normalized rat lung, Bento Soares"

Query Match 100.0%; Score 9; DB 21; Length 228;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 180 TTCTGAGA 188
|||||||
Cy 25 ttctcgaa 33

RESULT 58
LOCUS AA572725 233 bp mRNA EST 12-SEP-1997
DEFINITION n18a04.s1 NCI-CGAP_P1 Homo sapiens cDNA clone IMAGE:914094
similar to SW:510C_RABIT P24480 CALGIZARIN ;, mRNA sequence.
ACCESSION AA572725
NID G2347254
VERSION AA572725.1 GI:2347254
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 233)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Dec 30, 1996 this sequence version replaced gi:1529567.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
DNA sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
CDNA distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Insert Length: 552 Std Error: 0.00
Seq primer: -40m13 fwd. ET from AmerSham
High quality sequence stop: 97.
Location/Qualifiers
1..233

/organism="Homo sapiens"
/note="Vector: pAMP10; Site_1: NotI, Site_2: EcorI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors. 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into pAMP10 by the DGE-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Kitzman."

FEATURES
SOURCE
/clone="IMAGE:914094"
/clone_lib="NCI-CGAP_P1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"

BASE COUNT 57 a 51 c 56 g 69 t
ORIGIN
Query Match 100.0%; Score 9; DB 15; Length 233;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 TTCTGAGAA 22
|||||||
OY 25 ttctgagaa 33

RESULT 59 AI000683 233 bp mRNA EST 27-AUG-1998

LOCUS 0564f02.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1610139 3',
DEFINITION mRNA sequence.

ACCESSION AI000683

NID 93191237

VERSION AI000683.1 GI:3191237

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 233)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jan 19, 1998 this sequence version replaced gi:2045877.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 296 Std Error: 0.00

Seq primer: -40m13 fwd. ET from AmerSham.

FEATURES Location/Qualifiers

SOURCE

1. 233

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from pooled bulk

breast tumor tissue, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified pT73

vector. This library is the normalized version of

NCI_CGAP_Br1.1. Library was constructed by Bento Soares

and M. Fatima Bonaldo.

/db_xref="taxon:9606"

/clone="IMAGE:1610139"

/clone_lib="NCI_CGAP_Br2"

/sex="female, pooled"

/tissue_type="breast"

/lab_host="DH10B"

BASE COUNT 44 a 58 c 37 g 94 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 233;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 162 TTCTGAGAA 170

|||||||

Cp 33 ttctgagaa 25

RESULT 60 AA552472 234 bp mRNA

LOCUS

EST

11-AUG-1997

DEFINITION n190h01.s1 NCI_CGAP_Pr11 Homo sapiens cDNA clone IMAGE:999793
Similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AA552472

NID 92322726

VERSION AA552472.1 GI:2322726

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 234)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Sep 21, 1992 this sequence version replaced gi:276287.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Rodrigo F.

Chuaqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Krizman, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bdrrp/image/image.html

Seq primer: -40m13 fwd. ET from AmerSham.

FEATURES Location/Qualifiers

SOURCE

1. 234

/organism="Homo sapiens"

/note="Organ: prostate; Vector: pAMP10; mRNA made from

normal prostatic epithelial cells; cDNA made by oligo-dT

priming. Non-directionally cloned. Size-selected on

agarose gel, average insert size 600 bp. Library made by

D. Krizman, NIH."

/db_xref="taxon:9606"

/clone="IMAGE:999793"

/clone_lib="NCI_CGAP_Pr11"

/sex="male"

/tissue_type="normal prostatic epithelial cells"

/lab_host="DH10B"

BASE COUNT 53 a 59 c 53 g 69 t

ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 234;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 213 TTCTGAGAA 221

|||||||

OY 25 ttctgagaa 33

RESULT 61 AA939036 235 bp mRNA EST 07-JUL-1998

LOCUS 056g03.s1 Soares_NFL_T-GBC_S1 Homo sapiens cDNA clone

DEFINITION IMAGE:1580884 3', mRNA sequence.

ACCESSION AA939036

NID 93098949

VERSION AA939036.1 GI:3098949

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 235)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jan 19, 1998 this sequence version replaced g1:2285172.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1028 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 145.
Location/Qualifiers

FEATURES

1. .235

/organism="Homo sapiens"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP GCBI) were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 3,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1580884"
/clone_lib="Soares_NFL_T-GBC_S1"
/lab_host="DH10B"

BASE COUNT 52 a 64 c 51 g 68 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 235;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 192 TTCTGAGA 200
|||||||
Oy 25 ttctgaga 33

RESULT 62
LOCUS AA621837 235 bp mRNA EST 21-OCT-1997
DEFINITION n19e03.s1 NCI_CGAP_Thy1 Homo sapiens cDNA clone IMAGE:1144348,
mRNA sequence.
ACCESSION AA621837
NID 92525713
VERSION AA621837.1 GI:2525713
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 235)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Nov 29, 1993 this sequence version replaced g1:430445.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdnp/image/image.html

Insert Length: 541 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 227.
Location/Qualifiers

FEATURES

1. .235

/organism="Homo sapiens"
/note="Vector: PAMP10; mRNA made from invasive thyroid tumor; cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
/db_xref="taxon:9606"
/clone="IMAGE:1144348"
/clone_lib="NCI_CGAP_Thy1"
/tissue_type="thyroid"
/lab_host="DH10B"

BASE COUNT 92 a 38 c 57 g 48 t

ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 235;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 51 TTCTGAGA 59
|||||||
Cp 33 ttctgaga 25

RESULT 63
LOCUS AA712986 235 bp mRNA EST 24-DEC-1997
DEFINITION 32546 Lambda-PRL2 Arabidopsis thaliana cDNA clone 215M177, mRNA
sequence.
ACCESSION AA712986
NID 92722903
VERSION AA712986.1 GI:27222903
KEYWORDS EST.
SOURCE chae cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
rosidae; Capprales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 235)
AUTHORS Newman,T., deBruin,J.,F.J., Green,P., Keegstra,K., Kende,H.,
Mcintosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomaslow,M.,
Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95146729
COMMENT On Sep 12, 1996 this sequence version replaced g1:1397994.

CONTACT: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@dm.cl.msu.edu
Seq primer: T7 dye primer.
Location/Qualifiers

FEATURES

1. .235

/organism="Arabidopsis thaliana"
/strain="var columbiana"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

```

/db_xref="taxon:3702"
/map="22; xp22.2; 18q12.2-q21.31"
/clone="215M17T7"
/clone_id="Lambda-PRL2"
BASE COUNT      59 a      47 c      60 g      60 t      9 others
ORIGIN
Query Match      100.0%; Score 9; DB 17; Length 235;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 210 TTCGAGAA 218
|||||
Cp 33 ttccagaa 25

RESULT 64
LOCUS      AA627463      237 bp      mRNA      EST      16-OCT-1997
DEFINITION nq49904.s1 NCI-CGAP_Co10 Homo sapiens CDNA clone IMAGE:1147254 3',
ACCESSION  AA627463
VERSION     92539558
KEYWORDS    AA627463.1 GI:2539558
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 237)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Dec 18, 1996 this sequence version replaced gi:1734363.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 159.
Location/Qualifiers
1..237
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT713D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo (N-Soares4)."
/db_xref="taxon:9606"
/map="17q21"
/clone="IMAGE:1147254"
/clone_id="NCI-CGAP_Co10"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
BASE COUNT      48 a      64 c      76 g      49 t
ORIGIN
Query Match      100.0%; Score 9; DB 16; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 126 TTCGAGAA 134
|||||
Cy 25 ttccgagaa 33

RESULT 65
LOCUS      AA382444      237 bp      mRNA      EST      21-APR-1997
DEFINITION EST95699 Testis I Homo sapiens CDNA 3' end, mRNA sequence.
ACCESSION  AA382444
VERSION     92034805
KEYWORDS    AA382444.1 GI:2034805
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 237)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Mai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodok,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Welx,F.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Pennon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M., and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
            96026280
JOURNAL     On Sep 1, 1995 this sequence version replaced.
MEDLINE     Other_ESTs: EST95700 THC154392
COMMENT     The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlavet@igr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/ngi/ngi.html)
            Seq primer: M13-21.
Location/Qualifiers
1..237
/organism="Homo sapiens"
/note="Organ: testis; Vector: pBluescript SK-; Site:1;
EcoRI: Site:2; XhoI"
/db_xref="ATCC (lnhost):186776"
/db_xref="taxon:9606"
/map="5"
/clone_id="Testis I"
/sex="male"
/dev_stage="adult"
BASE COUNT      93 a      37 c      58 g      49 t
ORIGIN
Query Match      100.0%; Score 9; DB 12; Length 237;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 66
LOCUS AA568811 237 bp mRNA EST 22-AUG-1997
DEFINITION nm070403.s1 NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1059461 3',
mRNA sequence.
ACCESSION AA568811
NID 92341865
VERSION AA568811.1 GI:2341865
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 237)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407131.

FEATURES
source
Location/Qualifiers
1..237
/organism="Homo sapiens"
/Note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Ronaldo (N-Soares4)."
/db_xref="taxon:9606"
/clone_id="NCI_CGAP_Co10"
/issue_type="colon tumor RER+"
/lab_host="DH10B"
BASE COUNT 66 a 67 c 55 g 49 t
ORIGIN
Query Match 100.0%; Score 9; DB 15; Length 237;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 182 TTCTCAGAA 190
|||||
CP 33 ttctcagaa 25

RESULT 67
LOCUS AA53186 237 bp mRNA EST 06-AUG-1997
DEFINITION 3', mRNA sequence.
ACCESSION AA53186
NID 91882893
VERSION AA53186.1 GI:1882893
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 237)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubouche, T., Gelsel, G., Jost, S.,
Kucaba, T., Lacey, M., Le, N., Lennon, G., Marra, M., Math, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404755.

FEATURES
source
Location/Qualifiers
1..237
/organism="Homo sapiens"
/Note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHVP, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
/db_xref="taxon:9606"
/clone_id="Soares_NhMPU_S1"
/issue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
BASE COUNT 80 a 46 c 47 g 64 t
ORIGIN
Query Match 100.0%; Score 9; DB 10; Length 237;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 164 TTCTCAGAA 172
|||||
CP 33 ttctcagaa 25

RESULT 68
LOCUS AA339525 239 bp mRNA EST 21-APR-1997
DEFINITION EST44617 Fetal brain I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA339525
NID 91991761
VERSION AA339525.1 GI:1991761
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 239)
AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulcher, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weissbrock, R.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Li, L.-I., Nguyen, D.T., Pelissier, S.M.,
Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelissier, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,

```

Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Fertile, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

On Sep 12, 1996 this sequence version replaced gi:1393608.

Other ESTs: THC124239

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .239

/organism="Homo sapiens"

/note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

/db_xref="ATCC (inhost):141139"

/db_xref="taxon:9606"

/clone_lib="Fetal brain I"

/sex="female"

/dev_stage="fetus, 24 wks"

/dev_stage="fetus, 24 wks"

BASE COUNT 58 a 61 c 72 g 45 t 3 others

ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 239;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 169 TTCTGAGAA 177

|||||

Cp 33 ttctcagaa 25

RESULT 69

LOCUS AA380520 240 bp mRNA EST 21-APR-1997

DEFINITION EST93514 Supt cells Homo sapiens CDNA 5' end similar to Pr22 protein, mRNA sequence.

ACCESSION AA380520

NID 92032860

VERSION AA380520.1 GI:2032860

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 240)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulmer, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-vel, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgeraid, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodex, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Fertile, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,

Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

On Sep 12, 1996 this sequence version replaced gi:1393608.

Other ESTs: THC170735

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .240

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

/db_xref="ATCC (inhost):184935"

/db_xref="taxon:9606"

/clone_lib="Supt cells"

/cell_type="T-lymphocyte"

BASE COUNT 73 a 49 c 50 g 68 t

ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 240;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 TTCTGAGAA 89

|||||

Cp 25 ttctcagaa 33

RESULT 70

LOCUS 242215 240 bp mRNA EST 08-NOV-1994

DEFINITION HSC0C0011 normalized infant brain CDNA Homo sapiens CDNA clone c-0ce01, mRNA sequence.

ACCESSION 242215

NID 9565319

VERSION 242215.1 GI:565319

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 240)

Auffray, C., Behar, G., Bols, F., Bouchier, C., da Silva, C., Devignes, M.D., Duprat, S., Houllatte, R., Juneau, M.N., Lamy, B., Lorenzo, F., Mitchell, H., Mariage-Sanson, R., Pietu, G., Pouliot, Y., Sebastiani-Rabakhtchis, C. and Tessier, A.

IMAG: molecular integration of the analysis of the human genome and its expression

C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

95277534

On Sep 21, 1992 this sequence version replaced gi:279052.

CONTACT: Genethon

Genexpress-Genethon

Genethon Centre de recherche sur le Genome Humain

1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE

Tel: 33160778698

Fax: 33160778698

Email: genexpress@genethon.fr

Single read.

Genexpress_library_id: C: Genexpress_sequence_id: y1c-0ce01
Seq primer: (-21)M3_universal.
Location/Qualifiers

FEATURES

Source

1..240

/organism="Homo sapiens"

/note="Organ: brain; Vector: lafmid BA; Site 1: HindIII; Site 2: NotI; sex=Female; dev_stage=3 months old; isolate=muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B.Souares, Psychiatry Dept. Columbia University, USA. Normalization_method: Bento Soares, P.N.A.S in press"

/db_xref="taxon:9606"

/clone="c-0ce01"

/clone_idb="normalized infant brain CDNA"

/sex="Female"

/tissue_type="total brain"

/dev_stage="3 months old"

/dev_stage="3 months old"

/dev_stage="3 months old"

/dev_stage="3 months old"

BASE COUNT 69 a 38 c 48 g 83 t 2 others

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 240;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 225 TTCGACAA 233

CP 33 ttctgagaa 25

LOCUS 71 241 bp mRNA EST 06-SEP-1995

DEFINITION EST79569 Human Placenta Homo sapiens cDNA similar to decay accelerating factor for complement (Ht:1358), mRNA sequence.

ACCESSION 29421

NID 9611519

VERSION 29421.1 GI:611519

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 241)

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuidner,R.A., Bult,C.J., Lee,N., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.W., Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.T., Georghagen,N.S.M., Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinklejr,P.S., Kelley,J.M., Klemm,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,D.F., Li,Y., Bedarrik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferlie,A., Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M., Gruber,P.J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haselstein,W.A., Fields,C., Fraser,C.M., and Venter,J.C.

Initial Assessment of Human Gene Diversity and Expression Patterns Based Upon 83 Million Basepairs of CDNA Sequence

Nature 377, 3-174 (1995)

Other-ESTs: THC23264

Contact: Venter, JC

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD 20878

Tel: 3018699056

Fax: 3018699423

Email: tdbinfo@tdb.tigr.org

For clone availability, additional sequence and expression

information related to this EST, please contact the TIGR Database

(tdbinfo@tdb.tigr.org)

Seq primer: M3 Reverse

Location/Qualifiers

1..241

/organism="Homo sapiens"

/note="Organ: placenta"

/db_xref="ATCC (lnhost):106211"

/db_xref="taxon:9606"

/clone_idb="Human Placenta"

/clone="c-0ce01"

/clone_idb="Human Placenta"

/clone="c-0ce01"

/clone_idb="Human Placenta"

/clone="c-0ce01"

/clone_idb="Human Placenta"

/clone="c-0ce01"

/clone_idb="Human Placenta"

Information related to this EST, please contact the TIGR Database
(tdbinfo@tdb.tigr.org)
Seq primer: M3 Reverse
Location/Qualifiers

FEATURES

Source

1..241

/organism="Homo sapiens"

/note="Organ: placenta"

/db_xref="ATCC (lnhost):106211"

/db_xref="taxon:9606"

BASE COUNT 41 a 32 c 76 g 90 t 2 others

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 241;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 131 TTCGACAA 139

OY 25 ttctgagaa 33

LOCUS 72 241 bp mRNA EST 10-JUN-1997

DEFINITION xz72f11.r1 Soares,total_fetus_Nb2HF8-9w Homo sapiens cDNA clone IMAGE:797037 5', mRNA sequence.

ACCESSION AA463582

NID 92188466

VERSION AA463582.1 GI:2188466

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 241)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Matra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

Wash-Merck EST Project 1997

Unpublished (1997)

On Jun 27, 1996 this sequence version replaced gi:1184735.

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

Unpublished (1997)

```

BASE COUNT      66 a      48 c      42 g      85 t
ORIGIN
Query Match      100.0%; Score 9; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 TTCGAGAA 227
|||||
Oy 25 ttctgagaa 33

RESULT 73
LOCUS      F12521      242 bp      mRNA      EST      14-MAR-1995
DEFINITION HSC3B8111 normalized infant brain cDNA Homo sapiens CDNA clone
            c-3bail, mRNA sequence.
ACCESSION  F12521
NID         9708518
VERSION     F12521.1 GI:708518
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 242)
            Aufferay,C., Behar,G., Bols,F., Bouchier,C., da Silva,C.,
            Devignes,M.D., Duprat,S., Houigatte,R., Jumeau,M.N., Lamy,B.,
            Lorenzo,F., Mitchell,H., Marlage-Samson,R., Pletu,G., Poullot,Y.,
            Sebastiani-Kabatchis,C. and Tessier,A.
            IMAGE: molecular integration of the analysis of the human genome
            and its expression
            C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
JOURNAL    MEDLINE
            95277534
COMMENT
            Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress library Idt: C; Genexpress_sequence_idt: Y1c-3bail
            Insert length: 1195 Std Error: 0.00
            Seq primer: (-21)M13-universal
            High quality sequence stop: 378.
            Location/Qualifiers
                1..242
                /organism="Homo sapiens"
                /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                Site_2: NotI; sex:Female; dev_stage=3 months old;
                isolate-muscular atrophy patient; tissue-type=total
                brain; total mRNA was oligo-(dT) primed and directionally
                cloned 5' -> 3' into the HindIII -> NotI sites of the
                lafmid BA vector. Clone library from B.Souares, Psychiatry
                Dept. Columbia University, USA. Normalization_method:
                Bento Soares, P.N.A.S. in press"
                /db_xref="taxon:9606"
                /map="947H01; 11; 11p12-11p14.1"
                /clone="c-3bail"
                /clone_lib="normalized infant brain CDNA"
                /sex="Female"
                /tissue_type="total brain"
                /dev_stage="3 months old"
BASE COUNT      60 a      49 c      64 g      66 t      3 others
ORIGIN
Query Match      100.0%; Score 9; DB 19; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 206 TTCGAGAA 214
|||||

```

```

Cp 33 ttctcagaa 25

RESULT 74
LOCUS      AA342105      243 bp      mRNA      EST      21-APR-1997
DEFINITION EST47561 Fetal spleen Homo sapiens CDNA 5' end similar to Golg4
            complex outer membrane protein, mRNA sequence.
ACCESSION  AA342105
NID         91994340
VERSION     AA342105.1 GI:1994340
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 243)
            Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
            Rapid cDNA sequencing (expressed sequence tags) from a
            directionally cloned human infant brain CDNA library
            Nature Genet. 4, 373-380 (1993)
JOURNAL    MEDLINE
            94004865
COMMENT
            On Sep 12, 1996 this sequence version replaced g1:1407411.
            Other ESTs: EST47560 THCI73043
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3016699056
            Fax: 3016699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/ngi/ngi.html)
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..243
                /organism="Homo sapiens"
                /note="Organ: spleen; Vector: pBluescript SK-; Site_1:
                EcoRI; Site_2: XhoI"
                /db_xref="AACC (Inhost):143690"
                /db_xref="taxon:9606"
                /clone_lib="Fetal spleen"
                /dev_stage="fetus"
BASE COUNT      102 a      40 c      54 g      47 t
ORIGIN
Query Match      100.0%; Score 9; DB 11; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 TTCGAGAA 25
|||||
Oy 25 ttctgagaa 33

RESULT 75
LOCUS      AA408476      243 bp      mRNA      EST      26-AUG-1998
DEFINITION EST02957 Mouse 7.5 dpc embryo ectoplacental cone CDNA library Mus
            musculus CDNA clone C0028F01 3' similar to R.rattus mRNA for
            ribosomal protein L11, score = 968, mRNA sequence.
ACCESSION  AA408476
NID         92067872
VERSION     AA408476.1 GI:2067872
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 243)
            Ko,M.S.H., Threat,T.A., Horton,J.H., Wang,X., Cui,X., Wang,X.,
            Pryor,E., Patis,J., Wells-Smith,J., Fujiwara,H., Yotsunoto,S. and
            Nakashima,H.
            Systematic analyses of mouse genes expressed in embryo implantation

```

JOURNAL site
COMMENT Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1395293.

Contact: Ko MSH
Center for Molecular Medicine and Genetics
Wayne State University
5047 Guillen Mall, Detroit, MI 48202
Tel: 3135776708
Fax: 3135776200
Email: mskocemb.biosci.wayne.edu
R. ratus mRNA for ribosomal protein L11, score = 968
Seq primer: M13 Forward.

FEATURES
Source

1. 243
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: embryo; Vector: pSPORT1 (Life Technologies);
Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
ectoplacental cone of 7.5-dpc embryos. The
double-stranded cDNA was synthesized from total RNAs with
an Oligo(dT) primer. The library was constructed by Minoru
S. H. Ko."
/db_xref="ATCC (host):1364379"
/db_xref="taxon:10090"
/clone_lib="Mouse 7.5 dpc embryo ectoplacental cone cDNA
library"
/sex="unknown"
/tissue_type="ectoplacental cone"
/dev_stage="embryonic day 7.5 postconception"
/lab_host="DH10B"

BASE COUNT 48 a 63 c 56 g 75 t 1 others
ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,236+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCTGAGAA 13
|||||
0y 25 ttctcgaga 33

RESULT 76
LOCUS AA759834 243 bp mRNA EST 23-JAN-1998
DEFINITION v81d04.t1 Barstead mouse irradiated colon MRLB7 Mus musculus cDNA
clone IMAGE:1177543 5', mRNA sequence.

ACCESSION AA759834
NID 92807628
VERSION AA759834.1 GI:2807628

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 243)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Marrin, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Giselle, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 9, 1998 this sequence version replaced gi:936878.

Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:635391
Seq primer: -28m13 rev2 ET from Amerham.

FEATURES
Source

1. 243
/organism="Mus musculus"
/strain="FVB/N"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I - oligo(dT) primer
[5'-TGTTAGCATCTGAGTGGAGCGGCCCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors [AATCGGATCCTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified p773
vector. Library constructed by Bob Barstead."
/db_xref="taxon:10090"
/map="21"
/clone="IMAGE:1177543"
/clone_lib="Barstead mouse irradiated colon MRLB7"
/dev_stage="8 weeks"
/lab_host="DH10B"

BASE COUNT 93 a 53 c 43 g 54 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,236+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 167 TTCTGAGAA 175
|||||
Cp 33 ttctcgaga 25

RESULT 77
LOCUS AA809659 245 bp mRNA EST 12-FEB-1998
DEFINITION nk6d07.s1 NCI-CGAP_C03 Homo sapiens cDNA clone IMAGE:1028653 3',
mRNA sequence.

ACCESSION AA809659
NID 92879065
VERSION AA809659.1 GI:2879065

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 245)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Jan 19, 1998 this sequence version replaced gi:2287056.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amerham.
Location/Qualifiers
1. 245
/organism="Homo sapiens"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization.

/db_xref="taxon:9606"
/map="13"
/clone="IMAGE:1028653"
/clone_1b="NCI_CGAP-C03"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"

BASE COUNT 64 a 72 c 60 g 49 t
ORIGIN
Query Match 100.0%; Score 9; DB 18; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 159 TTCTGAGA 167
|||||
Cp 33 ttctcagaa 25

RESULT 78
LOCUS D47177 246 bp mRNA EST 02-AUG-1995
DEFINITION R1CS12347A Rice green shoot Oryza sativa cDNA, mRNA sequence.
ACCESSION D47177
NID 9700886
VERSION D47177.1 GI:700886
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.

REFERENCE 1 (bases 1 to 246)
AUTHORS Sasaki,T., Miyao,A. and Yamamoto,K.
TITLE Rice cDNA from callus 1995
JOURNAL Unpublished (1995)
COMMENT

Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program
2-1-2 Kannondai, Tsukuba
Ibaraki,
Japan 305
Tel: 0298-38-7441
Fax: 0298-38-7468
Email: tsasaki@agr.affrc.go.jp
Insert Length: 373 Std Error: 0.00
High quality sequence stop: 327.
Location/Qualifiers

FEATURES
source 1. 246
/organism="Oryza sativa"
/strain="Nipponbare"
/note="Green shoot (8 days old)"
/db_xref="taxon:4530"
/clone_1b="Rice green shoot"
BASE COUNT 72 a 74 g 64 t 3 others
ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 138 TTCTGAGA 146
|||||
Oy 25 ttctcagaa 33

RESULT 79

LOCUS T30847 246 bp mRNA EST 06-SEP-1995
DEFINITION EST23709 Human Bone Homo sapiens cDNA 5' end similar to None, mRNA sequence.
ACCESSION T30847
NID 9612945
VERSION T30847.1 GI:612945
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,O.A., Brandon,R.C., Chin,M.W.,
Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.M.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,P.S.,
Kelley,J.M., Kline,K.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M.,
Merrick,J.M., Moreno-Parlanques,R.F., McDonald,L.A., Nguyen,D.T.,
Pellegriano,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A.,
Coleman,T.A., Collins,E.J., Dinko,D., Feng,P., Ferris,A.,
Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H.,
L.H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
Hasseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 83 Million Basepairs of cDNA Sequence
Nature 377, 3-174 (1995)
JOURNAL MEDLINE
MEDLINE 96026280
COMMENT

Other-ESTs: TNC931
Contact: Venter, JC
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: tdbinfo@tdb.tigr.org
For clone availability, addtional sequence and expression
information related to this EST, please contact the TIGR Database
(tdbinfo@tdb.tigr.org)
Seq primer: M13 Reverse.

FEATURES
source 1. 246
/organism="Homo sapiens"
/db_xref="ATCC (inhost):100804"
/db_xref="taxon:9606"
/clone_1b="Human Bone"
/tissue_type="bone"

BASE COUNT 72 a 38 c 40 g 96 t
ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 237 TTCTGAGA 245
|||||
Cp 33 ttctcagaa 25

RESULT 80
LOCUS AA822615 246 bp mRNA EST 17-FEB-1998
DEFINITION v552612.t1 Striatagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:1149910 5', mRNA sequence.
ACCESSION AA822615
NID 92892483
VERSION AA822615.1 GI:2892483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 246)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The Washu-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 9, 1998 this sequence version replaced gi:936891.

Contact: Marra M/Mouse EST Project
 Washu-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:623118
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 151.

FEATURES
 SOURCE

1. .246
 /organism="Mus musculus"
 /note="Organ: blood; Vector: pBluescript SK-; Site:1; EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTT TTT TTT TTT TTT 3'"/
 /db_xref="taxon:10090"
 /map="21"
 /clone_1lb="Stratagene mouse Tcell 937311"
 /tissue_type="Tcell"
 /dev_stage="M30 CD4+ cells"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 72 a 68 c 50 g 56 t
 ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 246;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 46 TTCTCAGAA 54
 111111111
 Cp 33 ttctcagaa 25

RESULT 81
 LOCUS AA646505 247 bp mRNA EST 28-OCT-1997
 DEFINITION vni2105.r1 stratagene mouse heart (#937316) Mus musculus cDNA clone
 IMAGE:1021017 5', mRNA sequence.
 ACCESSION AA646505
 NID 92572934
 VERSION AA646505.1 GI:2572934
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 247)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 TITLE The Washu-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Sep 1, 1995 this sequence version replaced.

Contact: Marra M/Mouse EST Project
 Washu-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:571793
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 238.

FEATURES
 SOURCE

1. .247
 /organism="Mus musculus"
 /strain="NIH/Swiss"
 /note="Organ: heart; Vector: pBluescript SK-; Site:1; EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTT TTT TTT TTT TTT 3'"/
 /db_xref="taxon:10090"
 /map="3"
 /clone_1lb="Stratagene mouse heart (#937316)"
 /sex="pooled"
 /tissue_type="heart"
 /dev_stage="13 day embryos"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 58 a 67 c 75 g 47 t
 ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 247;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 117 TTCTCAGAA 125
 111111111
 QY 25 ttctcagaa 33

RESULT 82
 LOCUS AA565709 247 bp mRNA EST 08-SEP-1997
 DEFINITION nk47h12.s1 NCI-CCAP-GC2 Homo sapiens cDNA clone IMAGE:1016711 3',
 mRNA sequence.
 ACCESSION AA565709
 NID 92337348
 VERSION AA565709.1 GI:2337348
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 247)
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1393648.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.
 CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2354 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 215.
 Location/Qualifiers

FEATURES
 source
 1. .247
 /organism="Homo sapiens"
 /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Cloned unidirectionally. Primer: Oligo dt. Bulk
 germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCAG 3'
 3' adaptor sequence: 5' CTCACGTTTTTTTTTTTTTTT 3'
 Average insert size: 1.2 kb."
 /db_xref="taxon:9606"
 /clone_1fb="NCI_CGAP_GC2"
 /clone_1fb="NCI_CGAP_GC2"
 /tissue_type="germ cell tumor"
 /lab_host="SOLR (kanamycin resistant)"
 BASE COUNT 102 a 31 c 47 g 67 t
 ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 247;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 70 TTCGAGAA 78
 |||||
 Oy 25 ttctgagaa 33

RESULT 83
 LOCUS AA344114 249 bp mRNA EST 21-APR-1997
 DEFINITION EST49984 Gall bladder I Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA344114
 NID g1996351
 VERSION AA344114.1 GI:1996351
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 249)
 Bult.C.J., Lee.N.H., Kirkness.E.F., Wellstock.K.G., Gocayne.J.D.,
 White.O., Sutton.G., Blake.J.A., Brandon.R.C., Man-wei.C.,
 Clayton.R.A., Cline.T.R., Cotton.M.D., Earle-Hughes.E., Fine.L.D.,
 Fitzgerald.L.M., Fitzhugh.W.M., Fritchman.V.L., Geoghegan.N.S.,
 Glodex.A., Gnehm.C.L., Hanna.M.C., Hedblom.E., Hinkle.P.S.Jr.,
 Kelley.J.M., Kelley.J.C., Liu.L.-I., Marmaros.S.M., Merrick.J.M.,
 Moreno-Palanges.R.F., McDonald.L.A., Nguyen.D.T., Pellegrino.S.M.,
 Phillips.C.A., Ryder.S.E., Scott.J.L., Saudex.D.M., Shirley.R.,
 Small.R.V., Spriggs.T.A., Uterback.T.R., Weidman.J.F., Li.Y.,
 Bednarek.D.P., Cao.L., Cepeda.M.A., Coleman.T.A., Collins.E.J.,
 Dinko.D., Feng.D.-F., Ferris.A., Fischer.C., Hastings.G.A.,
 He.W.W., Hu.J.S., Greene.J.M., Gruber.J., Hudson.P., Kim.A.K.,
 Kosak.D.L., Kunsch.C., Hungjun.J., Li.H., Weissner.P.S., Olsen.H.,
 Raymond.L., Wei.Y.F., Wang.J., Xu.C., Yu.G.L., Ruben.S.M.,
 Dillion.P.J., Fannon.M.R., Rosen.C.A., Haseltine.W.A., Fields.C.,
 Fraser.C.M. and Venter.J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 On Apr 14, 1993 this sequence version replaced gi:716386.

TITLE
 JOURNAL
 MEDLINE
 COMMENT

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavage@igf.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/ngl/ngl.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES
 source
 1. .249
 /organism="Homo sapiens"
 /note="Organ: gall bladder; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"
 /db_xref="ATCC (Inositol):145639"
 /db_xref="taxon:9606"
 /map="942a12; 11; 11q24.3-11q25"
 /clone_1fb="Gall bladder I"
 /sex="female"
 /dev_stage="adult, 25 yrs"
 BASE COUNT 74 a 76 c 43 g 55 t 1 others
 ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 249;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 97 TTCGAGAA 105
 |||||
 Cp 33 ttctcagaa 25

RESULT 84
 LOCUS AA872080 250 bp mRNA EST 17-MAR-1998
 DEFINITION o112a11.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1476284 3',
 mRNA sequence.
 ACCESSION AA872080
 NID g2968258
 VERSION AA872080.1 GI:2968258
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 250)
 NCI_CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Nov 29, 1993 this sequence version replaced gi:501924.

JOURNAL
 AUTHORS
 TITLE
 COMMENT

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILM at:
 www-bio.llnl.gov/bdpr/image/image.html
 Seg primer: -40m13 fwd. ET from Amersham.
 Location/Qualifiers

FEATURES
 source
 1. .250
 /organism="Homo sapiens"
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 Oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pTZ19
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /map="22"
 /clone_1fb="NCI_CGAP_GC4"
 /tissue_type="pooled germ cell tumors"

BASE COUNT 70 a 48 c 57 g 75 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 20; Length 250;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 196 TTCTGAGAA 204
 |||||
 OY 25 ttctgagaa 33

RESULT 85
 LOCUS AA640401 250 bp mRNA EST 27-OCT-1997
 DEFINITION nt99e04.s1 NCI_CGAP_AlV1 Homo sapiens CDNA clone IMAGE:120654,
 mRNA sequence.
 ACCESSION AA640401
 NID 92565651
 VERSION AA640401.1 GI:2565651
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Euthera; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 21, 1992 this sequence version replaced g1:276411.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 CDNA Library Preparation: David B. Kitzman, Ph.D.
 CDNA Library Arrayed by: Greg Lennan, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 203.
 Location/Qualifiers
 1..250
 /organism="Homo sapiens"
 /note="Vector: PAMPI0; RNA made from alveolar
 rhadomyosarcoma, CDNA made by oligo-dT priming.
 Non-directionally cloned. Size selected on agarose gel,
 average insert size 600 bp. Reference: Kitzman et al.
 (1996) Cancer Research 56:5380-5383."
 /db_xref="taxon:9606"
 /clone="IMAGE:120654"
 /clone_1lb="NCI_CGAP_AlV1"
 /tissue_type="alveolar rhadomyosarcoma"
 /lab_host="DH10B"
 /lab_host="DH10B"

BASE COUNT 58 a 68 c 64 g 60 t
 ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 250;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 43 TTCTGAGAA 51
 |||||
 OY 25 ttctgagaa 33

RESULT 86
 LOCUS AA343598 252 bp mRNA EST 21-APR-1997

DEFINITION EST49407 Gall bladder I Homo sapiens CDNA 5' end, mRNA sequence.
 ACCESSION AA343598
 NID 91995837
 VERSION AA343598.1 GI:1995837
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Euthera; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 252)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bull,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'hai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.,
 Glodex,A., Gnehm,C.L., Hanna,M.C., Heblum,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Fertile,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Frazer,C.M. and Venter,J.C.
 TITLE Initial assessment of human gene diversity and expression patterns
 JOURNAL based upon 83 million nucleotides of CDNA sequence
 MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
 COMMENT On Apr 14, 1993 this sequence version replaced g1:693051.
 96026280
 Other_ESTS: TNC137512
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (http://www.tigr.org/tcbl/hgi/hgi.html)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..252
 /organism="Homo sapiens"
 /note="Organ: gall bladder; Vector: pBluescript SK-;
 Site_1: ECORI; Site_2: XhoI"
 /db_xref="ATCC (lnhost):145109"
 /db_xref="taxon:9606"
 /clone_1lb="Gall bladder I"
 /sex="female"
 /dev_stage="adult, 25 yrs"
 /dev_stage="adult, 25 yrs"

BASE COUNT 73 a 55 c 67 g 55 t
 ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 252;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 86 TTCTGAGAA 94
 |||||
 OY 25 ttctgagaa 33

RESULT 87
 LOCUS AA905871 252 bp mRNA EST 09-JUN-1998
 DEFINITION 0182a08.s1 Soares,NFL.T.GBC.S1 Homo sapiens CDNA clone
 IMAGE:1504790 3', mRNA sequence.
 ACCESSION AA905871
 NID 93040994
 VERSION AA905871.1 GI:3040994

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 252)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:716837.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 726 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 66.
Location/Qualifiers
1..252
/organism="Homo sapiens"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NDHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/clone="IMAGE:1504790"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/lab="52 c 49 g 85 t
BASE COUNT 66 a 52 c 49 g 85 t
ORIGIN
Query Match 100.0%; Score 9; DB 20; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 TTCTGAGAA 128
|||||||
Qy 25 ttctgagaa 33
RESULT 88
LOCUS AA743482 254 bp mRNA EST 27-JAN-1998
DEFINITION n721h01.s1 NCI CGAP GCB1 Homo sapiens cDNA clone IMAGE:1272433 3'
similar to contigs Ll.b2 Ll repetitive element; mRNA sequence.
ACCESSION AA743482
NID 92782988
VERSION AA743482.1 GI:2782988
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 254)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 9, 1998 this sequence version replaced gi:934519.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov

Tissue Procurement: Louis M. Staedt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D., Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
www-bio.llnl.gov/bdrp/image/image.html
Insert length: 1153 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1..254
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staedt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-CTTACCAATCGAATGGAGCGCGCCGCTCATTTTCTTTTCTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/db_xref="IMAGE:1272433"
/clone="NCI-CGAP-GCB1"
/clone_lib="germinal center B cell"
/issue_type="germinal center B cell"
/lab_host="DH10B"
/lab="57 a 41 c 35 g 121 t
BASE COUNT 57 a 41 c 35 g 121 t
ORIGIN
Query Match 100.0%; Score 9; DB 17; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.22e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 50 TTCTGAGAA 58
|||||||
Qy 25 ttctgagaa 33
RESULT 89
LOCUS A1014786 254 bp mRNA EST 16-JUN-1998
DEFINITION ot79f02.s1 Soares total fetus ND2HF8_9w Homo sapiens cDNA clone
IMAGE:1623003 3', mRNA sequence.
ACCESSION A1014786
NID 93229122
VERSION A1014786.1 GI:3229122
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 254)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Feb 10, 1998 this sequence version replaced gi:2339903.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 119.
Location/Qualifiers

source

1. 254
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTACCAATCTGATGAGGAGCGCGCTAATTTTGTGTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

/db_xref="taxon:9606"
/clone_id="Soares-total_fetus_Nb2Hf8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"

BASE COUNT 109 a 49 c 54 g 42 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 89 TTCTCAGAA 97
|||||
33 ttctcagaa 25

RESULT 90
LOCUS AA527319 255 bp mRNA EST
DEFINITION n36905.s1 NCI-CGAP_C03 Homo sapiens cDNA clone IMAGE:936849 3'
Similar to gb:X5755.tma2 SPLICING FACTOR SC35 (HUMAN), mRNA
sequence.
ACCESSION AA527319
NID 92269388
VERSION AA527319.1 GI:2269388
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 255)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1316069.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 352 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
Location/Qualifiers
1. 255
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library went through one round of

normalization."

/db_xref="taxon:9606"
/clone_id="IMAGE:936849"
/clone_id="NCI-CGAP_C03"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"

BASE COUNT 95 a 39 c 40 g 81 t

ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 255;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTCTCAGAA 9
|||||
25 ttctcagaa 33

RESULT 91
LOCUS AA593154 256 bp mRNA EST
DEFINITION n0607.s1 NCI-CGAP_P4.1 Homo sapiens cDNA clone IMAGE:1076940.
mRNA sequence.
ACCESSION AA593154
NID 92408916
VERSION AA593154.1 GI:2408916
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 256)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:785764.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuang,
M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arraying by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 379 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
Location/Qualifiers
1. 256
/organism="Homo sapiens"
/note="Organ: prostate; Vector: pAMP10; mRNA made from
prostatic intraepithelial neoplasia (high-grade), cDNA
made by oligo-dT priming. Non-directionally cloned.
Size selected on agarose gel, average insert size 600 bp.
CDNA Library Preparation: David B. Kitzman, Ph.D.
References: Kitzman et al. (1996) Cancer Research
56:5380-5383. CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing
Center"
/db_xref="taxon:9606"
/map="1/921"
/clone_id="IMAGE:1076940"
/clone_id="NCI-CGAP_P4.1"
/sex="male"
/tissue_type="prostatic intraepithelial neoplasia - high
grade"
/lab_host="DH10B"

BASE COUNT 91 a 34 c 41 g 90 t

ORIGIN 93

Query Match 100.0%; Score 9; DB 15; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 50 TTCTGAGA 58
|||||
0y 25 ttctgagaa 33

RESULT 92
LOCUS AA429636 257 bp mRNA EST 16-OCT-1997
DEFINITION zv74h05.s1 Soares_testis_NHT Homo sapiens CDNA clone IMAGE:781977
ACCESSION AA429636
NID g2112737
VERSION AA429636.1 GI:2112737
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 257)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenger, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.
Washu-Merck EST Project 1997
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1407467.
COMMENT

FEATURES
SOURCE
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 152.
Location/Qualifiers
1. 257
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from mRNA obtained from Clontech
laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5',
TGTTCACATCTGAGTGGAGCGCCGCCCAATTTTITTTTITTTT 3']
Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot₁ and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:781977"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
BASE COUNT 88 a 44 c 39 g 86 t
ORIGIN

Query Match 100.0%; Score 9; DB 13; Length 257;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 209 TTCTGAGA 217
|||||
0y 25 ttctgagaa 33

RESULT 93

LOCUS AA991905 259 bp mRNA EST 03-JUN-1998
DEFINITION os76g01.s1 NCI_CGAP_GC2 Homo sapiens CDNA clone IMAGE:1611312 3',
mRNA sequence.
ACCESSION AA991905
NID g3178787
VERSION AA991905.1 GI:3178787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 259)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1407356.
COMMENT

FEATURES
SOURCE
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Kitzman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 243.
Location/Qualifiers
1. 259
/organism="Homo sapiens"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: oligo dT. Bulk
gem cell tumor. 5' adaptor sequence: 5' GAAATTCGACGAG 3'
3' adaptor sequence: 5' CTCGACTTTTITTTTITTTTITTTT 3'
Average insert size: 1.2 kb."
/db_xref="taxon:9606"
/clone="IMAGE:1611312"
/clone_lib="NCI_CGAP_GC2"
/tissue_type="gem cell tumor"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 73 a 58 c 62 g 66 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.22e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 TTCTGAGA 38
|||||
Cp 33 ttctgagaa 25

RESULT 94

LOCUS AA496492 261 bp mRNA EST 11-AUG-1997
DEFINITION zv36D07.r1 Soares ovary tumor NDHOT Homo sapiens CDNA clone
IMAGE:755701 5', mRNA sequence.
ACCESSION AA496492
NID g2229813
VERSION AA496492.1 GI:2229813
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 261)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,

Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Scepcoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
Washu-Merck EST Project 1997
Unpublished (1997)
On Sep 21, 1992 this sequence version replaced g1:276370.

FEATURES
source
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.lbnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham.
Location/Qualifiers
1..261
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT773D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - 0190(dt) primer [5'
TGTACCAATCTGAGGAGCGCGCGCGGTTTGTGTGTGTGT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT773 vector
(Pharmacia). Library constructed by Bento Soares and
M.Felima Bonaldi.
/db_xref="taxon:9606"
/clone="IMAGE:755701"
/clone_lib="Soares ovary tumor NDHOT"
/sex="Female"
/tissue="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/lab_host="71 c 52 g 68 t

BASE COUNT 70 a 52 g 68 t
ORIGIN
Query Match 100.0%; Score 9; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 TTCTGAGAA 154
|||||||
Oy 25 ttctgagaa 33
RESULT 95
LOCUS AA357180 261 bp mRNA EST 21-APR-1997
DEFINITION EST65872 Jurkat T-cells I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA357180
NID G2009510
VERSION AA357180.1 GI:2009510
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 261)
AUTHORS Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE Rapid cDNA sequencing (expressed sequence tags) from a
directionally cloned human infant brain cDNA library
JOURNAL Nature Genet. 4, 373-380 (1993)
MEDLINE 94004965
COMMENT On Sep 12, 1996 this sequence version replaced g1:1288030.
Other_ESTs: THC81248
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression

Information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1..261
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
/db_xref="ATCC (Inhost):161572"
/db_xref="taxon:9606"
/clone_lib="Jurkat T-cells I"
/cell_type="T-lymphocyte"

BASE COUNT 85 a 37 c 31 g 107 t 1 others
ORIGIN
Query Match 100.0%; Score 9; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 158 TTCTGAGAA 166
|||||||
Oy 25 ttctgagaa 33
RESULT 96
LOCUS AA383812 262 bp mRNA EST 21-APR-1997
DEFINITION EST97241 Testis I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA383812
NID G2036130
VERSION AA383812.1 GI:2036130
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 262)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-vel,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.N., Fitzhugh,W.M., Fitchman,J.L., Geoghegan,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Margaros,S.M., Merrick,J.M.,
Moreno-Palenzuela,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudik,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
JOURNAL based upon 83 million nucleotides of cDNA sequence
MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
COMMENT Other_ESTs: EST183917 THC82058
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse
Location/Qualifiers
1..262
/organism="Homo sapiens"

FEATURES
source
Location/Qualifiers
1..262
/organism="Homo sapiens"

/note="Organ: testis; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="ATCC (lhost):187991"
/db_xref="taxon:9606"
/map="9"
/clone_lib="Testis I"
/sex="male"
/dev_stage="adult"
BASE COUNT 54 a 58 c 55 g 92 t 3 others
ORIGIN

Query Match 100.0%; Score 9; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 107 TTCGAGAA 115
|||||
Oy 25 ttctgagaa 33

RESULT 97
LOCUS AA721968 262 bp mRNA EST 02-JAN-1998
DEFINITION zh17b04.s1 Soares_pineal_gland.N3HPG Homo sapiens CDNA clone
IMAGE:412303 3', mRNA sequence.
AA721968
ACCESSION g2739675
NID AA721968.1 GI:2739675
VERSION EST.
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 262)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kilman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
On Sep 19, 1997 this sequence version replaced gi:1520524.

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jlnl.gov) for further information.
Seq primer: -40m13 fwd ER from Amersham.

FEATURES
Location/Qualifiers
1..262
/organism="Homo sapiens"
/note="Organ: pineal gland; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5', TGTTACCAATCTGAGTGGAGCGCCGCTTTTCTTTTCTTTTCTT
3'], double-stranded cDNA was size selected, ligated to
Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT73
vector (Pharmacia). Library constructed by Bento Soares
and M. Fatima Bonaldi.
/db_xref="GDB:130397"
/db_xref="taxon:9606"
/map="20"
/clone="IMAGE:412303"
/clone_lib="Soares_pineal_gland.N3HPG"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 67 a 62 c 64 g 69 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 164 TTCGAGAA 172
|||||
Oy 25 ttctgagaa 33

RESULT 99
LOCUS AA197928 263 bp mRNA EST 12-MAR-1997

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 88 TTCGAGAA 96
|||||
Oy 25 ttctgagaa 33

RESULT 98
LOCUS R03503 263 bp mRNA EST 31-MAR-1995
DEFINITION PK08C03.r1 Kuwabara Mixed stage C. briggsae Caenorhabditis briggsae
CDNA, mRNA sequence.
R03503
ACCESSION g753239
NID R03503.1 GI:753239
VERSION EST.
KEYWORDS
SOURCE
ORGANISM Caenorhabditis briggsae.
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditiia; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Peloderiinae; Caenorhabditis.
1 (bases 1 to 263)
Hillier, L., Chapell, B., Chisoe, S., Clark, N., Couch, J.,
Dubuque, T., Hawkins, M., Holman, M., Hultman, M., Kucaba, T.,
Kuwabara, P., Le, M., Mardis, E., Marra, M., Parsons, J., Rifkin, L.,
Rohlfing, T., Tan, F., Trevisakis, E., Waterston, R., Wohldmann, P. and
Wilson, R.
Washington University Caenorhabditis briggsae EST project
Unpublished (1995)
Other-ESTs: PK08C03.s1
Contact: Marra MA
Washington University Genome Sequencing Center
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1455
Fax: 314 286 1810
Email: mmarra@watson.wustl.edu
PCR_F: TGTAACACGACGCGCCGTCAGCAGTTTCAGCCTGG
PCR_R: CAGGAACAGATATGACCTTATGAGTATTCCTCAGCGTA
Source: Washington University Genome Sequencing Center
PCR amplified DNA is available from Washington University Genome
Sequencing Center. Aliquots of the library may be requested from P.
Kuwabara (pek@mc-lmb.cam.ac.uk).
Insert Length: 494 Std Error: 0.00
Seq primer: Commercially available M13 reverse dye primer
High quality sequence stop: 278.

TITLE
JOURNAL
COMMENT

Contact: Marra MA
Washington University Genome Sequencing Center
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1455
Fax: 314 286 1810
Email: mmarra@watson.wustl.edu
PCR_F: TGTAACACGACGCGCCGTCAGCAGTTTCAGCCTGG
PCR_R: CAGGAACAGATATGACCTTATGAGTATTCCTCAGCGTA
Source: Washington University Genome Sequencing Center
PCR amplified DNA is available from Washington University Genome
Sequencing Center. Aliquots of the library may be requested from P.
Kuwabara (pek@mc-lmb.cam.ac.uk).
Insert Length: 494 Std Error: 0.00
Seq primer: Commercially available M13 reverse dye primer
High quality sequence stop: 278.

FEATURES
Location/Qualifiers
1..263
/organism="Caenorhabditis briggsae"
/strain="G16 Gujarati"
/note="Vector: lambda g110; Site_1: EcoRI; Site_2: EcoRI;
Stage: mixed, Sex: hermaphrodite. Library construction:
First strand oligo(dT) primed. Second strand was as per
Gubler/Hofman. Ligated to EcoRI adapters. Library is
non-directional. Library is non-normalized. Library
constructed by P.E. Kuwabara. Additional details on
construction of the library are described in P.E.
Kuwabara and S. Shah, NAR 22: 4414 - 4418 (1994). Adaptor
sequence: GAATTC CGTTCGCTGCG"
/db_xref="taxon:6238"
/clone_lib="Kuwabara Mixed stage C. briggsae"
/clone="lambda g110"
BASE COUNT 82 a 56 c 54 g 69 t 2 others
ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 263;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 164 TTCGAGAA 172
|||||
Oy 25 ttctgagaa 33

RESULT 99
LOCUS AA197928 263 bp mRNA EST 12-MAR-1997

DEFINITION mw10b04.r1 Guay/Woodford Belter mouse kidney day 0 Mus musculus cDNA
clone IMAGE:654607 5' similar to gb:M2324 AMINOPEPTIDASE N
(HUMAN); mRNA sequence.

ACCESSION AA197928
NID g1793684
VERSION AA197928.1 GI:1793684
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 263)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Seisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The Mashu-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced g1:1394324.

CONTACT: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:400455
Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES
source Location/Qualifiers
1..263
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: kidney; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned undirectionally. Primer:
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector:
-5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' Library provided
Lisa Guay-Woodford."
/db_xref="taxon:10090"
/clone="IMAGE:654607"
/clone_1lb="Guay/Woodford Belter mouse kidney day 0"
/tissue_type="kidney"
/dev_stage="newborn (day 0)"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 64 a 67 c 76 g 56 t
ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 263;
Best Local Similarity 100.0%; Pred. NO. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 240 TTCTGAGAA 248
|||||
Oy 25 ttctgagaa 33

RESULT 100
LOCUS AA525063 263 bp mRNA EST 05-AUG-1997
DEFINITION nh32903.s1 NCI-CGAP_Pt3 Homo sapiens cDNA clone IMAGE:954100
similar to gb:S41458 ROD CGMP-SPECIFIC 3',5'-CYCLIC
PHOSPHODIESTERASE BETA-SUBUNIT (HUMAN); mRNA sequence.
ACCESSION AA525063
NID g2265991
VERSION AA525063.1 GI:2265991
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 263)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced g1:785493.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Kitzman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Insert length: 493 Std Error: 0.00
Seq primer: -40ml3 fwd ET from Amersham
High quality sequence stop: 88.

FEATURES
source Location/Qualifiers
1..263
/organism="Homo sapiens"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into PAMP10 by the UDG-cloning method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Kitzman."
/db_xref="taxon:9606"
/clone="IMAGE:954100"
/clone_1lb="NCI-CGAP_Pt3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
BASE COUNT 82 a 50 c 53 g 78 t
ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. NO. 2.22e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 159 TTCTCAGAA 167
|||||
Cp 33 ttctcagaa 25

RESULT 101
LOCUS D33548 264 bp mRNA EST 08-AUG-1994
DEFINITION CELK032E9R yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone yK32e9 3', mRNA sequence.
ACCESSION D33548
NID g524467
VERSION D33548.1 GI:524467
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditidae;
Rhabditina; Rhabditioidea; Rhabditidae; Pterodermidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 264)
AUTHORS Kohara, Y., Mitsui, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and
Tabara, H.
TITLE Toward an expression map of the C. elegans genome
JOURNAL Unpublished (1994)
COMMENT On Sep 21, 1992 this sequence version replaced g1:279326.

Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.

FEATURES
source
Location/Qualifiers

1. .264
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/note="dev_stage=varied; sex=Hermaphrodite male;
tissue_type=whole animal"
/db_xref="taxon:6239"
/clone="yk3269"
/clone_lib="Yuji Kohara unpublished cDNA"
BASE COUNT 100 a 33 c 76 g 55 t
ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 TTCTCAGAA 41
Cp 33 ttctcagaa 25

RESULT 102
LOCUS 239705 264 bp mRNA EST 07-NOV-1994
DEFINITION HSC1H102 normalized infant brain cDNA Homo sapiens CDNA clone
c-1hd10 3, mRNA sequence.
ACCESSION 239705
NID 9565500
VERSION 239705.1 GI:565500
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C.,
Devignes, M.D., Duprat, S., Houligatte, R., Juneau, M.N., Lamy, B.,
Loienzo, F., Mitchell, H., Mariège-Samson, R., Pietu, G., Pouliot, Y.,
Sebastien-Kabatchis, C. and Tessier, A.

TITLE
JOURNAL
MEDLINE
COMMENT
IMAG: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr

Single read. 28 T removed at sequence 5' end
Genexpress_library_idt: C; Genexpress_sequence_idt: a1c-1hd10
Seq primer: (-21)M13 universal.
Location/Qualifiers

1. .264
/organism="Homo sapiens"
/note="Organ: brain; Vector: lafmid BA; Site:1: HindIII;
Site:2: NotI; sex=Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization method:
Bento Soares, P.N.A.S in press"
/db_xref="taxon:9606"

/map="1"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
BASE COUNT 97 a 45 c 44 g 70 t 8 others
ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 170 TTCTCAGAA 178
Cp 33 ttctcagaa 25

RESULT 103
LOCUS AA266518 264 bp mRNA EST 20-MAR-1997
DEFINITION m67a11.r1 Soares mouse lymph node NbMLN Mus musculus CDNA clone
IMAGE:718460 5' similar to gp:X36740 RAS-RELATED PROTEIN RAB-11
(HUMAN);, mRNA sequence.
ACCESSION AA266518
NID 91902616
VERSION AA266518.1 GI:1902616
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE
JOURNAL
COMMENT
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Sep 21, 1992 this sequence version replaced gi:276285.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:443956

Trace considered overall poor quality
Putative full length read
Vector to vector length is
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.

Location/Qualifiers

1. .264
/organism="Mus musculus"
/strain="57BL/6J"
/note="Organ: lymph node; Vector: p17T3D-Pac (Pharmacia)
with a modified polylinker; Site:1: Not I; Site:2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5',
TGTACCAATCTGATGAGGAGGAGGCGCCGACATCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p17T3 vector. RNA
provided by Dr. Bertrand Jordan. Library constructed and
normalized by Bento Soares and M.Fatima Bonaldo."
/db_xref="taxon:10090"
/clone="IMAGE:718460"
/clone_lib="Soares mouse lymph node NbMLN"

```
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT      76 a      61 c      50 g      77 t

ORIGIN

Query Match      100.0%; Score 9; DB 10; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      6 TTCTCAGAA 14
        |||||||
Cp      33 ttctcagaa 25

RESULT 104
LOCUS      AA587547      264 bp      mRNA      EST      25-SEP-1997
DEFINITION nd44a12.s1 NCI_CGAP_A1 Homo sapiens cDNA clone IMAGE:803134 3'
            similar to contains L1.t1 L1 repetitive element ;, mRNA sequence.
ACCESSION      AA587547
NID      92397222
VERSION      AA587547.1 GI:2397222
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 264)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
TITLE      Unpublished (1997)
JOURNAL      On Apr 14, 1993 this sequence version replaced gi:837673.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Lee Helman, Ph.D., Michael Emmert-Buck, M.D.,
Ph.D.
CDNA Library Preparation: Life Technologies Inc., David Kitzman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bdrp/image/image.html

Insert Length: 1397 Std Error: 0.00
Seq primer: -40ml3 fwd. Ef from Amersham
High quality sequence stop: 210.
Location/Qualifiers

I. 264
/organism="Homo sapiens"
/note="Vector: PCMV-SPORT2; Site_1: SalI; Site_2: NotI;
Cloned unidirectionally. Primer: Oligo dt."
/db_xref="GDB:6029006"
/db_xref="taxon:9606"
/clone_image="IMAGE:803134"
/clone_lib="NCI_CGAP_A1"
/tissue_type="bulk alveolar tumor"
/lab_host="DH10B"

BASE COUNT      95 a      34 c      49 g      86 t

ORIGIN

Query Match      100.0%; Score 9; DB 15; Length 264;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      217 TTCTCAGAA 225
        |||||||
Cp      33 ttctcagaa 25
```

```
RESULT 105
LOCUS      AA280530      265 bp      mRNA      EST      14-AUG-1997
DEFINITION zt1f12.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712847 3'
            similar to TR:G998901 G998901 PDCD2-PROGRAMMED CELL DEATH-2/RP8
            HOMOLOG . ;, mRNA sequence.
ACCESSION      AA280530
NID      91924766
VERSION      AA280530.1 GI:1924766
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 265)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
TITLE      Unpublished (1997)
JOURNAL      On Sep 12, 1996 this sequence version replaced gi:1402085.
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Insert Length: 846 Std Error: 0.00
Seq primer: -41ml3 fwd. Ef from Amersham
High quality sequence stop: 195.
Location/Qualifiers

I. 265
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - Oligo(dT) primer
5'-ggttaccatctgagtgagcgccgccctcattttttttttttttt-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_image="IMAGE:712847"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT      45 a      84 c      66 g      69 t      1 others

ORIGIN

Query Match      100.0%; Score 9; DB 10; Length 265;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      177 TTCTCAGAA 185
        |||||||
Cp      25 ttctcagaa 33

RESULT 106
LOCUS      AA603628      266 bp      mRNA      EST      08-OCT-1997
DEFINITION np19g01.s1 NCI_CGAP_Br3 Homo sapiens cDNA clone IMAGE:1116816 3'
            mRNA sequence.
ACCESSION      AA603628
NID      92437489
VERSION      AA603628.1 GI:2437489
KEYWORDS      EST.
SOURCE      human.
```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 266)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL

On Sep 12, 1996 this sequence version replaced gi:1397717.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES

Source

Insert Length: 1507 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 261.

Location/Qualifiers

1..266
/organism="Homo sapiens"
/note="Organ: breast; Vector: Bluescript SK-; Site: 1;
EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer:
Oligo dt. Ductal breast tumor. 5' adaptor sequence: 5'
GAATTCGCGACGAG 3' 3' adaptor sequence: 5'
CTCAGCTTTTCTTTTCTTTTCTTTT 3' Average insert size: 0.9 kb."
/db_xref="taxon:9606"
/map="21"
/clone="IMAGE:1116816"
/clone_1lb="NCI_CGAP_Br3"
/sex="female"
/tissue_type="breast tumor"
/lab_host="SOLR (Kanamycin resistant)"
BASE COUNT 86 a 48 c 38 g 93 t 1 others

BASE COUNT

Query Match 100.0%; Score 9; DB 15; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 148 TTCTGAGAA 156
|||||||
Oy 25 ttctgagaa 33

RESULT 107

LOCUS AA904030 266 bp mRNA EST 21-APR-1998
DEFINITION aa94e11.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1417388 3',
mRNA sequence.

ACCESSION

AA904030
g3039153
AA904030.1 GI:3039153

VERSION

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 266)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
On Jan 24, 1995 this sequence version replaced gi:634370.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 827 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 73.

FEATURES

Source

1..266
/organism="Homo sapiens"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoma, and was then primed with a
Not I - oligo(dt) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1417388"
/clone_1lb="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
BASE COUNT 77 a 50 c 47 g 92 t

Query Match 100.0%; Score 9; DB 20; Length 266;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 180 TTCTGAGAA 188
|||||||
Oy 25 ttctgagaa 33

RESULT 108

LOCUS AA241715 266 bp mRNA EST 22-SEP-1997
DEFINITION MB25J8C9T3 Brugia malayi second stage larvae JHU96SL-Bml2 Brugia
malayi cDNA clone L2S38C9 5', mRNA sequence.

ACCESSION

AA241715
g1870981
AA241715.1 GI:1870981

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1996)
On Jun 18, 1996 this sequence version replaced gi:1366626.

Contact: Blaxter ML

Institute of Cell, Animal and Population Biology

University of Edinburgh

Marworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9

3JT, UK.

Tel: +44 131 650 6760

Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk

The ABI trace of this sequence can be viewed at

<http://www.sanger.ac.uk/brugia/L2S/MB25J8C9T3.html> This is the

full sequence of the cDNA clone. The polyA tail has been clipped

and is excluded from this sequence

Seq primer: T3.

FEATURES
source

Location/Qualifiers
1. 266
/organism="Brugia malayi"
/strain="TRS Labs"
/note="Vector: LambdaZapII; Site:1: Eco R I (5' end); Site:2: Xho I (3' end); Brugia malayi is a lymphatic filarial nematode parasite of humans. Full length cDNA was prepared by long-range RT-PCR from mRNA from L2 larvae of the human filarial nematode parasite Brugia malayi using nematode spliced leader (SL, 5'end) and oligo-d(T) (3' end) primers. The library had an unamplified titre of ~1 x 10E6 per ml and ~95% of clones have inserts (mean length ~900 bp). The library is available from The Filarial Genome Project Resource Center: contact Dr. S.A. Williams, Clark Science Center, Smith College, Northampton, MA 01063 usa phone +1 413 585 3826 fax +1 413 585 3786 email genome@smith.edu"
/db.xref="taxon:6279"
/clone.lib="Brugia malayi second stage larvae JH096SL-Bml2"
/sex="mixed"
/dev_stage="mosquito derived, second stage larvae (L2)"
/lab_host="E. coli XL1-Blue MRF-"
BASE COUNT 75 a 37 c 51 g 98 t 5 others
ORIGIN

Query Match 100.0%; Score 9; DB 10; Length 266;
Best local similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 TTCTGAGAA 113
|||||
CP 33 ttctcgaa 25

RESULT 109
LOCUS AA197883 270 bp mRNA EST 12-MAR-1997
DEFINITION mw04d11.t1 GuayWoodford Beler mouse kidney day 0 Mus musculus CDNA
clone IMAGE:654045 5' similar to TR:G406113 G406113 PROTEIN KINASE
1.; mRNA sequence.
ACCESSION AA197883
NID G1793500
VERSION AA197883.1 GI:1793500
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 270)
AUTHORS Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394984.

FEATURES
source
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:399893
Seq primer: -28ml3 rev1 ET from AmerSham
High quality sequence stop: 229.
Location/Qualifiers
1. 270

/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: Kidney; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: 019g dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3' Library provided Lisa Guay-Woodford."
/db.xref="taxon:10090"
/clone.lib="GuayWoodford Beler mouse kidney day 0"
/clone.lib="GuayWoodford Beler mouse kidney day 0"
/issue_type="Kidney"
/dev_stage="newborn (day 0)"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 63 a 62 c 74 g 71 t
ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 270;
Best local similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 154 TTCTGAGAA 162
|||||
QY 25 ttctgagaa 33

RESULT 110
LOCUS AA276995 270 bp mRNA EST 01-APR-1997
DEFINITION va23h12.t1 GuayWoodford Beler mouse kidney day 7 Mus musculus CDNA
clone IMAGE:732455 5', mRNA sequence.
ACCESSION AA276995
NID G1919631
VERSION AA276995.1 GI:1919631
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 270)
AUTHORS Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1325182.

FEATURES
source
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:449503
High quality sequence stop: 234.
Location/Qualifiers
1. 270
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: Kidney; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: 019g dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3' Library provided by Lisa Guay-Woodford."
/db.xref="taxon:10090"
/clone.lib="GuayWoodford Beler mouse kidney day 7"
/issue_type="Kidney"

/dev_stage="juvenile (7 days old)"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 78 a 58 c 76 g 58 t

ORIGIN

Query Match 100.0%; Score 9; DB 10; Length 270;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 TTCGAGAA 159
|||||
Cp 33 ttctcagaa 25

RESULT 111
LOCUS AA642354 271 bp mRNA EST 13-NOV-1997
DEFINITION ns29e08.s1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1185062 3',
ACCESSION AA642354
NID 92567572
VERSION AA642354.1 GI:2567572
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 271)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407082.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
unknown library type
Insert Length: 1426 Std Error: 0.00
Seq Primer: -40ml3 fwd: ET from Amersham.
Location/Qualifiers
1. .271
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human consillar cells enriched for
germinal center B cells by flow sorting (CD20+, 19D-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBPR). CDNA synthesis was
primed with a Not I - oligo(dt) primer
[5'-TGTTACCAATCTGAGGAGGCGCCCTCTTTTCTTTTCTTTT-
3']. Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalisation, and was
constructed by Bento Soares and M. Fatima Boudado."
/db_xref="taxon:9606"
/clone_image="1185062"
/clone_id="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 65 a 71 c 61 g 74 t

ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 TTCGAGAA 13
|||||
Cp 25 ttctcagaa 33

RESULT 112

LOCUS AA497742 271 bp mRNA EST 01-JUL-1997
DEFINITION v172c09.r1 Striatagene mouse testis (#937308) Mus musculus CDNA
clone IMAGE:917776 5', mRNA sequence.
ACCESSION AA497742
NID 92232765
VERSION AA497742.1 GI:2232765
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 271)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Watson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Dec 18, 1996 this sequence version replaced gi:1734306.

CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:529992
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 252.
Location/Qualifiers
1. .271
/organism="Mus musculus"
/strain="Inbred CD-1"
/note="Organ: testis; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GAATTCGGCAGCG 3' -3' adaptor
sequence: 5' CTCGAGCTTTTCTTTTCTTTTCTTTT 3."
/db_xref="taxon:10090"
/clone_image="917776"
/clone_id="Striatagene mouse testis (#937308)"
/sex="males"
/tissue_type="testis"
/dev_stage="10-12 week old"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 71 a 84 c 63 g 53 t

ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 61 TTCGAGAA 69
|||||
Cp 33 ttctcagaa 25

RESULT 113
LOCUS AA182421 271 bp mRNA EST 10-MAR-1998
DEFINITION zps56f02.r1 Striatagene NT2 neuronal precursor 937230 Homo sapiens
CDNA clone IMAGE:613467 5', mRNA sequence.
ACCESSION AA182421
NID 91766304
VERSION AA182421.1 GI:1766304
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 271)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST project
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced g1:716991.

FEATURES
 SOURCE Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 3799 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 127.
 Location/Qualifiers
 1..271

BASE COUNT 83 a 49 c 81 g 54 t 4 others
 ORIGIN
 Query Match 100.0%; Score 9; DB 9; Length 271;
 Best Local Similarity 100.0%; Pred. No. 2,236+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 85 TTCGAGAA 93
 Cp 33 ttctcgaag 25

RESULT 114
 LOCUS AA640402 272 bp mRNA EST 27-OCT-1997
 DEFINITION nc99e05.s1 NCI-CGAP_Alv1 Homo sapiens CDNA clone IMAGE:1206556.
 ACCESSION AA640402
 NID 92565652
 VERSION AA640402.1 GI:2565652
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 272)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 21, 1992 this sequence version replaced g1:276412.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 DNA distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbrrp/image/image.html

FEATURES
 SOURCE Seg primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 143.
 Location/Qualifiers
 1..272

FEATURES
 SOURCE /organism="Homo sapiens"
 /note="Vector: PAMPI0; mRNA made from alveolar rhabdomyosarcoma, CDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel." average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
 /db_xref="taxon:9606"
 /clone="IMAGE:1206656"
 /clone_lib="NCI-CGAP_Alv1"
 /tissue_type="alveolar rhabdomyosarcoma"
 /lab_host="DH10B"

BASE COUNT 61 a 79 c 67 g 65 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 16; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2,236+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 43 TTCGAGAA 51
 Cp 25 ttctcgaag 33

RESULT 115
 LOCUS AA990483 273 bp mRNA EST 02-JUN-1998
 DEFINITION ua60h12.s1 Soares 2NDMT Mus musculus CDNA clone IMAGE:1361927 3' similar to SW:G10_XENIA P12805 G10 PROTEIN. ;, mRNA sequence.
 ACCESSION AA990483
 NID 93175847
 VERSION AA990483.1 GI:3175847
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 273)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMT Mouse EST Project
 JOURNAL Unpublished (1998)
 COMMENT On Jan 19, 1998 this sequence version replaced g1:2286559.

Contact: Marra M/Mouse EST Project
 WashU-HMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:699967
 Possible reversed clone: similarity on wrong strand
 Possible reversed clone: polyT not found
 Seq primer: -40m13 fwd. ET from Amersham.
 Location/Qualifiers
 1..273

FEATURES
 SOURCE /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCATCTGAGTCGAGCGCCGCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

/db_xref="taxon:10090"
/map="x"
/clone="IMAGE:1361927"
/clone_lib="Soares 2dbm"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT 62 a 77 c 73 g 61 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 273;
Best Local Similarity 100.0%; Pzed. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 236 TTCTCAGAA 244
|||||
ttctcagaa 25

RESULT 116
LOCUS Aa684949 274 bp mRNA EST 03-APR-1998
DEFINITION EST105836 Rat PC-12 cells, untreated Rattus sp. cDNA clone RPCAT58
5' end similar to Elongation factor 1 gamma, mRNA sequence.
ACCESSION Aa684949
NID 92671547
VERSION Aa684949.1 GI:2671547
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 274)
Lee,N.H., Weinstein,K.G., Kirnness,E.F., Earle-Hughes,J.A., Fuldner,R.A., Marmaras,S., Glodex,A., Geacayne,J.D., Adams,M.D., Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment
Proc Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
95396786
On Nov 29, 1993 this sequence version replaced gi:502312.
Other ESTs: TC45098
Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.

FEATURES
Source
Location/Qualifiers
1..274
/organism="Rattus sp."
/note="(Vector: Bluescript SK-; Site.1: EcoRI; Site.2: XhoI; poly(A)+ RNA was purified from untreated PC12 cells cultured for 9 days. cDNA was constructed using an oligo-dT primer and directionally cloned using the lambda ZAP II Vector Kit by Stratagene"
/db_xref="ATCC (inhost):2000826"
/db_xref="taxon:10118"
/map="1"
/clone="RPCAT58"
/clone_lib="Rat PC-12 cells, untreated"

BASE COUNT	72 a	67 c	76 g	58 t	1 others
ORIGIN					
Query Match	100.0%	Score 9;	DB 16;	Length 274;	
Best Local Similarity	100.0%	Pred. No. 2,236+03;			
Matches	9;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Db	3 TTCTCAGAA 11				
CP	33 ttctcagaa 25				
RESULT 117					
LOCUS	AA422037	275 bp	mRNA	EST	16-OCT-1997
DEFINITION	zv8db06.s1 Soares ovary tumor NBHOT Homo sapiens CDNA clone				
IMAGE	754931 3', mRNA sequence.				
ACCESSION	AA422037				
NID	92100853				
VERSION	AA422037.1	GI:2100853			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 275) Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kuwab, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellendberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.				
TITLE	Mashu-Merck EST Project 1997				
JOURNAL	Unpublished (1997)				
COMMENT	On Apr 14, 1993 this sequence version replaced gi:692564.				
	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LINT ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -41m3 fwd. ET from Amersham High quality sequence stop: 223.				
FEATURES	Location/Qualifiers				
SOURCE	1. 275 /organism="Homo sapiens" /note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAAGTGGACGCGCGGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Felina Bonaldo."				
	/db_xref="taxon:9606" /clone="IMAGE:754931" /clone_id="Soares Ovary tumor NBHOT" /sex="Female" /issue_type="Ovarian tumor" /lab_host="DH10B (ampicillin resistant)"				
BASE COUNT	92 a	51 c	43 g	89 t	
ORIGIN					
Query Match	100.0%	Score 9;	DB 13;	Length 275;	
Best Local Similarity	100.0%	Pred. No. 2,236+03;			
Matches	9;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Db	148 TTCTCAGAA 156				
ORIGIN	25 ttctcagaa 33				

RESULT 118
 LOCUS AA302718 275 bp mRNA EST 18-APR-1997
 DEFINITION EST10242 Adipose tissue, white I Homo sapiens cDNA 3' end, mRNA
 sequence.
 ACCESSION AA302718
 NID 91955049
 VERSION AA302718.1 GI:1955049
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 275)
 Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
 Rapid cDNA sequencing (expressed sequence tags) from a
 directionally cloned human infant brain cDNA library
 Nature Genet. 4, 373-380 (1993)
 JOURNAL 94004965
 MEDLINE
 COMMENT On Apr 14, 1993 this sequence version replaced gi:635723.
 Other_ESTs: THC177532
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: akkerlavet@igf.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tcdb/hgi/hgi.html>)
 Seq primer: M13-21.
 FEATURES
 source
 1..275
 /organism="Homo sapiens"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 /db_xref="taxon:9606"
 /map="13"
 /clone_lib="Adipose tissue, white I"
 /tissue_type="adipose tissue, white"
 /dev_stage="adult"
 /base_count="88 a 74 c 55 g 52 t 6 others"
 BASE COUNT 88 a 74 c 55 g 52 t 6 others
 ORIGIN
 Query Match 100.0%; Score 9; DB 11; Length 275;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 156 TTCTGAGAA 164
 |||||||
 Cp 33 ttctcagaa 25
 RESULT 119
 LOCUS AA282274 276 bp mRNA EST 13-AUG-1997
 DEFINITION zt132405.t1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:712952 5',
 mRNA sequence.
 ACCESSION AA282274
 NID 91925388
 VERSION AA282274.1 GI:1925388
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 276)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT On Nov 29, 1993 this sequence version replaced gi:430552.

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Putative full length read
 The vector to vector length is 277
 Seq primer: -28m3 rev2 ET from Amersham.
 FEATURES
 source
 1..276
 /organism="Homo sapiens"
 /note="Vector: p1773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+, IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 (5'-TGTTACCACTGAGAGGAGCGCCGCTATTTTCTTTTCTTTT-
 3'). Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p1773 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /map="21"
 /clone_image="712952"
 /clone_lib="NCI-CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 BASE COUNT 77 a 63 c 75 g 61 t
 ORIGIN
 Query Match 100.0%; Score 9; DB 10; Length 276;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 175 TTCTGAGAA 183
 |||||||
 Oy 25 ttctcagaa 33
 RESULT 120
 LOCUS F10065 277 bp mRNA EST 26-FEB-1995
 DEFINITION HSC3AB042 normalized infant brain cDNA Homo sapiens cDNA clone
 c-3ab04 3', mRNA sequence.
 ACCESSION F10065
 NID 9682774
 VERSION F10065.1 GI:682774
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 277)
 Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
 Devignes,M.D., Duprat,S., Houllgate,R., Jumeau,M.N., Lamy,B.,
 Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
 Sebastiani,Kabakchis,C. and Tessier,A.
 IMAGE: molecular integration of the analysis of the human genome
 and its expression
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 JOURNAL 95277534
 MEDLINE
 COMMENT On Sep 21, 1992 this sequence version replaced gi:278956.
 CONTACT: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr
 Single read. removed at sequence 5'end

Genexpress_library_idt: C: Genexpress_sequence_idt: a1c-3ab04
Seq primer: (-21)M13_universal
High quality sequence stop: 254
Location/Qualifiers
1. .277

FEATURES
source
/organism="Homo sapiens"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI; sex:Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soures, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
/db_xref="taxon:9606"
/clone_11b="3ab04"
/clone_11b="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
BASE COUNT 91 a 48 c 52 g 84 t 2 others
ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 28 TTCGTGAGA 36
|||||||
Oy 25 ttctgagaa 33

RESULT 121
LOCUS AA954232 278 bp mRNA EST 07-MAY-1998
DEFINITION o092b04.s1 NCI_CGAP_K1d5 Homo sapiens cDNA clone IMAGE:1573615 3',
mRNA sequence.
ACCESSION AA954232
NID g3117927
VERSION AA954232.1 GI:3117927
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 278)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2152365.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrp/image/image.html

FEATURES
source
Seg primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 277.
Location/Qualifiers
1. .278

/organism="Homo sapiens"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
ACTGTGAAGATTCGGCGCCGCAATATATTTTTTTTTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/map="21g"
/clone="IMAGE:1573615"
/clone_11b="NCI_CGAP_K1d5"
/tissue_type="2 pooled tumors (clear cell type)"
/lib_host="DH10b"
BASE COUNT 89 a 54 c 50 g 85 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 260 TTCGTGAGA 268
|||||||
Oy 25 ttctgagaa 33

RESULT 122
LOCUS AA938205 279 bp mRNA EST 13-MAY-1998
DEFINITION o010f10.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1340491 3',
similar to SW:YA94.SCHPO Q09783 HYPOTHETICAL 11.4 KD PROTEIN
C13G6.04 IN CHROMOSOME 1.; mRNA sequence.
ACCESSION AA938205
NID g3096316
VERSION AA938205.1 GI:3096316
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 279)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2152693.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 486 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 210.
Location/Qualifiers
1. .279

/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTACCAATTCGACGTGAGAGCGCCGCAATATATTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1340491"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT 69 a 73 c 57 g 79 t 1 others
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 84 TTCTGAGAA 92
|||||
OY 25 ttctgagaa 33

RESULT 123
LOCUS AA764915 279 bp mRNA EST 07-FEB-1998
DEFINITION n60b04.s1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1299823 3',
mRNA sequence.
ACCESSION AA764915
NID 92816153
VERSION AA764915.1 GI:2816153
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 279)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 17, 1998 this sequence version replaced gi:1900396.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 1655 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 239.
Location/Qualifiers
1. 279
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD+),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - oligo(dt) primer
[5'-TGTTACATCTGAGTGGAGCGCCCTCTTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1299823"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"

/lab_host="DH10B"
BASE COUNT 83 a 68 c 53 g 75 t
ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 TTCTGAGAA 34
|||||
Cp 33 ttctcagaa 25

RESULT 124
LOCUS AA953790 280 bp mRNA EST 07-JUL-1998
DEFINITION o038a11.s1 NCI_CGAP_Lus Homo sapiens CDNA clone IMAGE:1568444 3',
mRNA sequence.
ACCESSION AA953790
NID 9316708
VERSION AA953790.1 GI:3116708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 280)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 17, 1998 this sequence version replaced gi:2044305.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html

Insert Length: 457 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 272.
Location/Qualifiers
1. 280
/organism="Homo sapiens"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoma, and was then primed with a
Not I - oligo(dt) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/map="3"
/clone="IMAGE:1568444"
/clone_lib="NCI_CGAP_Lus"
/tissue_type="carcinoid"
/lab_host="DH10B"

BASE COUNT 79 a 68 c 70 g 63 t
ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 32 TTCTGAGAA 40
|||||

Cp 33 ttctcagaa 25

RESULT 125

LOCUS AA629000 280 bp mRNA EST 16-OCT-1997
 DEFINITION U73C09.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:743632
 3' mRNA sequence.

ACCESSION

AA629000
 NID 92541387
 VERSION AA629000.1 GI:2541387
 KEYWORDS EST.

SOURCE

ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 280)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
 Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
 Theising,B., White,T., Wylie,T., Waterston,R. and Wilson,R.
 WashU-NCI human EST Project
 Unpublished (1997)

TITLE JOURNAL
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1393489.

Contact: Wilson RK

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 266.

FEATURES

source

1..280

/organism="Homo sapiens"

/note="vector: pT7T3D-Pac (pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 laboratories, Inc., and primed with a Not I - 01190(dt)
 primer [5']

TGTACCAATCGAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo.

/db_xref="taxon:9606"
 /db_xref="taxon:9606"
 /clone="IMAGE:743632"
 /clone_11b="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"

/lab_host="DH10B"

BASE COUNT 81 a 67 c 66 g 66 t
 ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 280;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 246 TTCTCAGAA 254

Cp 33 ttctcagaa 25

RESULT 126

LOCUS AA589358 280 bp mRNA EST 16-SEP-1997
 DEFINITION v16a10.s1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
 IMAGE:75258 3', mRNA sequence.

ACCESSION

AA589358
 NID 92402738
 VERSION AA589358.1 GI:2402738

KEYWORDS

EST.
 house mouse.
 Mus musculus

SOURCE

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 280)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)

TITLE JOURNAL
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1404840.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:555986

Possible reversed clone: polyt not found
 Seq primer: -28m13 rev1 Et from Amersham
 High quality sequence stop: 189.

FEATURES

source

1..280

/organism="Mus musculus"

/strain="C57Bl/6"

/note="Organ: Skin; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 01190 dt. Whole skin from 11 week old C57Bl/6 female mice.
 Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5'
 adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
 sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'

/db_xref="taxon:10090"
 /clone="IMAGE:975258"
 /clone_11b="Stratagene mouse skin (#937313)"
 /sex="females"
 /tissue="types="whole skin"
 /dev_stage="11 weeks old"
 /lab_host="SOIR (kanamycin resistant)"

/lab_host="SOIR (kanamycin resistant)"

BASE COUNT 76 a 62 c 56 g 86 t
 ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 280;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 164 TTCTCAGAA 172

Cp 33 ttctcagaa 25

RESULT 127

LOCUS AA302617 282 bp mRNA EST 18-APR-1997
 DEFINITION E815792 Aorta endothelial cells Homo sapiens cDNA 5' end similar
 to similar to EBBR-associated protein, mRNA sequence.

ACCESSION

AA302617
 NID 91954948
 VERSION AA302617.1 GI:1954948
 KEYWORDS EST.

SOURCE

ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 282)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A.,
 Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,

Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Gloder, A., Gnehm, C.L., Hanna, M.C., Heddom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 Nature 377 (5547 Suppl), 3-174 (1995)

On Jun 18, 1996 this sequence version replaced g1:136592.
 Other_ESTS: THC188456
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3016699056
 Fax: 3016699423
 Email: atkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

Source
 1..282
 /organism="Homo sapiens"
 /note="Organ: aorta; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"
 /db_xref="ATCC (lnost):114794"
 /db_xref="taxon:9606"
 /clone_idb="aorta endothelial cells"
 /cell_type="endothelial cell"
 /dev_stage="adult"
 /dev_stage="adult"

BASE COUNT
 86 a 48 c 66 g 80 t 2 others

ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 282;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 184 TTCTGAGAA 192
 |||||
 25 ttctgagaa 33

RESULT 128
 LOCUS A1012299 283 bp mRNA EST 15-JUN-1998
 DEFINITION EST206750 Normalized rat placenta, Bento Soares Rattus sp. cDNA
 clone RPLAV23 3' end, mRNA sequence.

ACCESSION A1012299
 MID 93226131
 VERSION A1012299.1 GI:3226131
 KEYWORDS EST.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.

REFERENCE 1 (bases 1 to 283)
 Lee, N.H., Gloder, A., Chandra, I., Mason, T.M., Quackenbush, J., Kerlavage, A.R., and Adams, M.D.
 Rat Genome Project: Generation of a Rat EST (RESTR) Catalog & Rat Gene Index
 Unpublished (1998)
 On Jan 19, 1998 this sequence version replaced g1:2285611.

Contact: Lee, NH
 ARCC
 The Institute for Genomic Research
 9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@tigr.org
 Seq primer: M13-21

FEATURES

Source
 1..283
 /organism="Rattus sp."
 /note="Organ: placenta; Vector: pT733pac; Site_1: EcoRI; Site_2: NotI"
 /db_xref="taxon:10118"
 /clone="RPLAV23"
 /clone_idb="Normalized rat placenta, Bento Soares"
 /clone_idb="Normalized rat placenta, Bento Soares"

BASE COUNT 60 a 83 c 81 g 59 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 283;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 65 TTCTGAGAA 73
 |||||
 25 ttctgagaa 33

RESULT 129
 LOCUS AA652376 283 bp mRNA EST 13-NOV-1997
 DEFINITION ns52h09.s1 NCI_CGAP_P22 Homo sapiens cDNA clone IMAGE:1187297 3', mRNA sequence.

ACCESSION AA652376
 MID 92584028
 VERSION AA652376.1 GI:2584028
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 283)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 On Sep 12, 1996 this sequence version replaced g1:1397808.

JOURNAL

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html

Insert Length: 375 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

Source
 1..283
 /organism="Homo sapiens"
 /note="Organ: prostate; Vector: pT733-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT733 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo." /db_xref="taxon:9606"

```

/map="18"
/clone="IMAGE:1187297"
/clone_lib="NCI_CGAP_Pr22"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH108"

BASE COUNT      96 a      47 c      57 g      83 t

ORIGIN

Query Match      100.0%; Score 9; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      204 TTCTCAGAA 212
      111111111
      25 ttctcagaa 33

RESULT 130
LOCUS      AA296986      283 bp      mRNA      EST      18-APR-1997
DEFINITION EST112537 Adrenal gland tumor Homo sapiens cDNA 5' end, mRNA
sequence.
ACCESSION      AA296986
NID      91949317
VERSION      AA296986.1 GI:1949317
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 283)
AUTHORS      Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
      Bulic,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
      White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
      Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
      Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.,
      Glodet,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
      Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
      Moreno-Pedraza,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
      Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shiley,R.,
      Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
      Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
      Dimek,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
      He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
      Kozak,D.L., Kunsch,C., Hungjun,T., Li,H., Melsner,P.S., Olsen,H.,
      Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
      Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
      Fraser,C.M. and Venter,J.C.
      Initial assessment of human gene diversity and expression patterns
      based upon 83 million nucleotides of cDNA sequence
      Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL      Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE      96026280
COMMENT      On Sep 12, 1996 this sequence version replaced g1:1326766.
      Other ESTs: THC177442
      Contact: Kerlavage, AR
      Bioinformatics
      The Institute for Genomic Research
      9712 Medical Center Drive, Rockville, MD 20850 USA
      Tel: 3018699056
      Fax: 3018699423
      Email: arkerlav@tigr.org
      For clone availability, additional sequence and expression
      information related to this EST, please check the TIGR Human Gene
      Index (http://www.tigr.org/tdb/hgi/hgi.html)
      Seq primer: M13 Reverse.
      Location/Qualifiers
        1..283
        /organism="Homo sapiens"
        /note="Organ: adrenal gland; Vector: pBluescript SK-;
        Site_1: EcoRI; Site_2: XhoI"
        /db_xref="ATCC (inhost):1120252"
        /db_xref="taxon:9606"
        /clone_lib="Adrenal gland tumor"
FEATURES
SOURCE

```

```

/dev_stage="adult"

BASE COUNT      84 a      68 c      64 g      67 t

ORIGIN

Query Match      100.0%; Score 9; DB 10; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      149 TTCTCAGAA 157
      111111111
      33 ttctcagaa 25

RESULT 131
LOCUS      AA927636      284 bp      mRNA      EST      10-JUN-1998
DEFINITION oml19g12.s1 Soares_NFL_T-GBC_S1 Homo sapiens cDNA clone
IMAGE:1541542 3', mRNA sequence.
ACCESSION      AA927636
NID      63077057
VERSION      AA927636.1 GI:3077057
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 284)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
      Tumor Gene Index
      Unpublished (1997)
JOURNAL      On Jan 19, 1998 this sequence version replaced g1:2152808.
COMMENT      Contact: Robert Strausberg, Ph.D.
      Tel: (301) 496-1550
      Email: Robert.Strausberg@nih.gov
      This clone is available royalty-free through LNL; contact the
      IMAGE Consortium (info@image.llnl.gov) for further information.
      Insert length: 1144 Std Error: 0.00
      Seq primer: -40ml3 fwd. ET from Amersham
      High quality sequence stop: 265.
      Location/Qualifiers
        1..284
        /organism="Homo sapiens"
        /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
        a modified polylinker; Site_1: Not I; Site_2: Eco RI;
        Equal amounts of plasmid DNA from three normalized
        libraries (fetal lung NBH19W, testis NHT, and B-cell
        NCI-CGAP GCBI) were mixed, and ss circles were made in
        vitro. Following HAP purification, this DNA was used as
        tracer in a subtractive hybridization reaction. The driver
        was PCR-amplified cDNAs from pools of 5,000 clones made
        from the same 3 libraries. The pools consisted of
        1.M.A.G.E. clones 297480-302087, 682632-687239,
        726408-728711, and 729096-733399. Subtraction by Bento
        Soares and M. Fatima Bonaldo.
        /db_xref="taxon:9606"
        /clone="IMAGE:1541542"
        /clone_lib="Soares_NFL_T-GBC_S1"
        /lab_host="DH108"
BASE COUNT      98 a      60 c      50 g      76 t

ORIGIN

Query Match      100.0%; Score 9; DB 20; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      201 TTCTCAGAA 209
      111111111
      33 ttctcagaa 25

RESULT 132
LOCUS      T60003      284 bp      mRNA      EST      09-FEB-1995

```

DEFINITION yco1e11.r1 Strataene lung (#937210) Homo sapiens cDNA clone IMAGE:79436 5' similar to gb:U68060 DNA TOPOLISOMERASE II, beta ISOZYME (HUMAN); contains MER6 repetitive element ; mRNA sequence.

ACCESSION 160003

NID 9661840

VERSION T60003.1 GI:661840

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

Insert Size: 54

High quality sequence stops: 212 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 54 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 212.

Location/Qualifiers

1. 284

/organism="Homo sapiens"

/note="Organ: lung; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGGAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTCCTTTTTCCTTTT 3' "

/db_xref="db:483053"

/db_xref="taxon:9606"

/clone_image="IMAGE:79436"

/clone_lib="Strataene lung (#937210)"

/sex="male"

/dev_stage="72 years"

/lab_host="SOLR cells (kanamycin resistant)"

BASE COUNT 96 a 39 c 70 g 78 t 1 others

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 284;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 126 TTCTCAGAA 134
|||||

Cp 33 ttctcagaa 25

RESULT 133

LOCUS AA907008 285 bp mRNA EST 09-JUN-1998

DEFINITION OX98c12.s1 NCI-CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1522006 3' similar to gb:M25667 NEUROMODULIN (HUMAN); mRNA sequence.

ACCESSION AA907008

NID 93042468

VERSION AA907008.1 GI:3042468

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 285)

COMMENT NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT On Jan 14, 1998 this sequence version replaced gi:1797384.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/dbfp/image/image.html

Insert Length: 583 Std Error: 0.00

Seq primer: -40mJ fwd. ET from Amersham

High quality sequence stop: 87.

Location/Qualifiers

1. 285

/organism="Homo sapiens"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

/db_xref="taxon:9606"

/clone_image="IMAGE:1522006"

/clone_lib="NCI-CGAP_Lu5"

/clone_type="cathoid"

/lab_host="DH10B"

BASE COUNT 73 a 48 c 96 g 68 t

ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 285;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 TTCTCAGAA 227
|||||

Cp 25 ttctcagaa 33

RESULT 134

LOCUS AA429624 292 bp mRNA EST 16-OCT-1997

DEFINITION zw74g10.s1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:781986 3' mRNA sequence.

ACCESSION AA429624

NID 92112725

VERSION AA429624.1 GI:2112725

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-Merck EST Project

JOURNAL Unpublished (1997)

COMMENT On Sep 12, 1996 this sequence version replaced gi:1404989.

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
 Email: estewatson.vustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41ml3 fwd. ET from Amersham
 High quality sequence stop: 271.

FEATURES

source

Location/Qualifiers

1..292

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech.
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5']
 TGTACCAATCTGAAGTGAGCGCGCCCAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to cots, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:781986"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /lab_host="5' c 77 g 60 t

BASE COUNT

90 a 65 c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

/lab_host="5' c 77 g 60 t

Query Match 100.0%; Score 9; DB 13; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2,236+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 157 TTCTGAGAA 165
 |||||
 25 ttctgagaa 33

RESULT 135
 LOCUS AA297030 292 bp mRNA EST 18-APR-1997
 DEFINITION EST112879 Embryo, 9 week Homo sapiens cDNA 5' end similar to
 transmembrane protein (GB:U19878, mRNA sequence).
 ACCESSION AA297030
 NID G1949678
 VERSION AA297030.1 GI:1949678
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 292)
 Adams,M.D., Kerlavage,A.R., Felschmann,R.D., Fulmer,R.A.,
 Bull,C.J., Lee,N.H., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
 Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dinko,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M., and Venter,J.C.

TITLE

JOURNAL

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

MEDLINE

96026280

Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1294277.
 Other_ESTs: THC177990

FEATURES

source

Location/Qualifiers

1..292

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"
 /db_xref="ATCC (Inhost):136147"
 /db_xref="taxon:9606"
 /clone_lib="Embryo, 9 week"
 /dev_stage="embryo, 9 wks"
 /dev_stage="embryo, 9 wks"

BASE COUNT

90 a 45 c 80 g 75 t 2 others

/dev_stage="embryo, 9 wks"

/dev_stage="embryo, 9 wks"

/dev_stage="embryo, 9 wks"

/dev_stage="embryo, 9 wks"

/dev_stage="embryo, 9 wks"

/dev_stage="embryo, 9 wks"

Query Match 100.0%; Score 9; DB 10; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2,236+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 TTCTGAGAA 32
 |||||
 Cp 33 ttctgagaa 25

RESULT 136
 LOCUS AA344440 292 bp mRNA EST 21-APR-1997
 DEFINITION EST50332 Gall bladder I Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA344440
 NID G1996679
 VERSION AA344440.1 GI:1996679
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 292)
 Adams,M.D., Kerlavage,A.R., Felschmann,R.D., Fulmer,R.A.,
 Bull,C.J., Lee,N.H., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
 Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dinko,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
 He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M., and Venter,J.C.

TITLE

JOURNAL

Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

MEDLINE

96026280

Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT

On Apr 14, 1993 this sequence version replaced gi:716809.
 Other_ESTs: THC178808

MEDLINE

96026280

Nature 377 (6547 Suppl), 3-174 (1995)

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1294277.
 Other_ESTs: THC177990

On Sep 12, 1996 this sequence version replaced gi:1294277.
 Other_ESTs: THC177990

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 301699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/ngi.html>)
 Seq primer: M13 Reverse.

Email: arkerlavett@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

Source
 Location/Qualifiers
 1..292
 /organism="Homo sapiens"
 /note="Organ: gall bladder; Vector: pBluescript SK-;
 Site_1: EcorI; Site_2: XhoI"
 /db_xref="ATCC (inhost):145979"
 /db_xref="taxon:9606"
 /map="18 pl.1-q11.2"
 /clone_lib="gall bladder I"
 /sex="female"
 /dev_stage="adult, 25 yrs"
 BASE COUNT 100 a 54 c 49 g 87 t 2 others
 ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 292;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 TTCTGAGAA 64
 Cp 33 ttctcgagaa 25

RESULT 137
 LOCUS T30360 293 bp mRNA EST 06-SEP-1995
 DEFINITION EST851594 Human Endothelial cells Homo sapiens cDNA 5' end similar
 to similar to ribosomal protein L14, mRNA sequence.
 ACCESSION T30360
 NID 9612458
 VERSION 130360.1 GI:612458
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 293)
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bait,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
 Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 FitzGerald,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghegan,N.S.M.,
 Glodex,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,P.S.,
 Kelley,J.M., Klimke,K.M., Kelley,J.C., Liu,L.-I., Martiros,S.M.,
 Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T.,
 Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
 Snudde,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
 Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A.,
 Coleman,T.A., Collins,E.-J., Dinke,D., Feng,P., Fertle,A.,
 Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
 Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
 Misener,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
 Yu,G.-L., Ruben,S.M., Dillon,P.J., Pannon,M.R., Rosen,C.A.,
 Haseilime,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
 Initial Assessment of Human Gene Diversity and Expression Patterns
 Based Upon 83 Million Basepairs of cDNA Sequence
 Nature 377, 3-174 (1995)
 96026280

Other ESTs: THC12476

CONTACT: Venter, J.C.
 The Institute for Genomic Research
 932 Clapper Rd, Gaithersburg, MD 20878
 Tel: 3018699056
 Fax: 3018699423

Journal MEDLINE
 COMMENT
 Email: tdbinfo@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please contact the TIGR Database
 (tdbinfo@tigr.org)
 Seq primer: M13 Reverse.

FEATURES

Source
 Location/Qualifiers
 1..293
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):100073"
 /db_xref="taxon:9606"
 /clone_lib="Human Endothelial cells"
 /tissue_type="endothelial"
 BASE COUNT 73 a 72 c 79 g 69 t
 ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 293;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 188 TTCTGAGAA 196
 Oy 25 ttctgagaa 33

RESULT 138
 LOCUS AA369164 294 bp mRNA EST 21-APR-1997
 DEFINITION EST80533 Placenta II Homo sapiens cDNA 5' end similar to EST
 containing Alu repeat, mRNA sequence.
 ACCESSION AA369164
 NID 92021483
 VERSION AA369164.1 GI:2021483
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 294)
 Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
 Rapid cDNA sequencing (expressed sequence tags) from a
 directionally cloned human infant brain cDNA library
 Nature Genet. 4, 373-380 (1993)
 94004965
 JOURNAL MEDLINE
 COMMENT
 On Sep 12, 1996 this sequence version replaced gi:1393187.
 Other ESTs: THC12749
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlavett@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

Source
 Location/Qualifiers
 1..294
 /organism="Homo sapiens"
 /note="Organ: placenta; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"
 /db_xref="ATCC (inhost):173838"
 /db_xref="taxon:9606"
 /clone_lib="placenta II"
 /tissue_type="placenta"
 /dev_stage="fetus"
 BASE COUNT 66 a 73 c 70 g 82 t 3 others
 ORIGIN

Query Match 100.0%; Score 9; DB 12; Length 294;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 243 TTCTGAGAA 251
 Oy 25 ttctgagaa 33

RESULT 139

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Ilian Kirsch, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNB at:
www-bio.llnl.gov/bdnp/image/image.html
 Trace considered overall poor quality
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. .295

[illegible]

CONTACT: Maira M/Mouse ESI Project
WashU-HMMI Mouse ESI Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseesi@wustl.edu
This clone is available royalty-free through LINT : contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
WGI:904580
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 274.
Location/Qualifiers
1..295

```

/Note="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primer
with a Not I - Q1igo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pTR73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:10090"
/Clone="IMAGE:1382112"
/clone_lib="Soares mouse mammary gland NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
BASE COUNT      88 a      62 c      56 g      89 t
ORIGIN
Query Match      100.0%; Score 9; DB 21; Length 295;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 179 TTCGAGCA 187
      |||||
      |||||
QY 25 ttctgaa.33
RESULT 141
LOCUS M62263 295 bp mRNA EST 26-MAY-1992
DEFINITION EST00332 Hippocampus, Striatum (cat. #356205) Homo sapiens cDNA
clone HHCJ48, mRNA sequence.
ACCESSION M62263
NID G272508
VERSION M62263.1 GI:272508
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS Adams,M.D., Kelley,J.M., Gocayne,J.D., Dubnick,M.,
Polymeropoulos,M.H., Xiao,H., Merril,C.R., Wu,A., Olde,B.,
Moreno,R.F. and etal.
TITLE Complementary DNA sequencing: expressed sequence tags and human
genome project
JOURNAL Science 252, 1651-1656 (1991)
MEDLINE 91262645

```

COMMENT

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
Seq primer: M13 Forward.
Location/Qualifiers

FEATURES

source

BASE COUNT

95 a 62 c 75 g 63 t

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 295;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 134 TTCACGAA 142
|||||

Cp 33 ttccagaa 25

RESULT 142

LOCUS 239383 296 bp mRNA EST 31-OCT-1994
DEFINITION HSC18H012 normalized infant brain CDNA Homo sapiens cDNA clone
c-18h01 3', mRNA sequence.

ACCESSION

239383
239383
239383.1 GI:562575

VERSION 239383.1 GI:562575
KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

AufRAY,C., Behar,G., Bols,F., Bouchier,C., da Silva,C.,
Devignes,M.D., Duprat,S., Houligatte,R., Jumeau,M.N., Lamy,B.,
Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pletu,G., Pouliot,Y.,
Sebastiani-Kabakchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression

TITLE

C.R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534

JOURNAL

MEDLINE

COMMENT

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read. 18 T removed at sequence 5' end
Genexpress_library_id: C; Genexpress_sequence_id: alc-18h01
Seq primer: (-21)M13-universal.

FEATURES

source

1. .296
/organism="Homo sapiens"
/note="Organ: brain; Vector: lafmid BA; Site.1: HindIII;
Site.2: NotI; sex=Female; dev.stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dfr) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Souares, Psychiatry

Dept. Columbia University, USA. Normalization_method:

Bento Soares, P.N.A.S in press"

/db_xref="taxon:9606"

/map="14"

/clone="C-18h01"

/sex="Female"

/tissue_type="total brain"

/dev_stage="3 months old"

BASE COUNT

102 a 58 c 60 g 76 t

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 296;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 172 TTCACGAA 180
|||||

Cp 33 ttccagaa 25

RESULT 143

LOCUS AA905891 298 bp mRNA EST 08-APR-1998
DEFINITION c182f08.s1 Soares,NFL.T.GBC.S1 Homo sapiens cDNA clone
IMAGE:1504839 3' similar to contains Alu repetitive element; mRNA
sequence.

ACCESSION AA905891

NID 93041014

VERSION AA905891.1 GI:3041014

KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 298)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Apr 14, 1993 this sequence version replaced gi:716859.

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m3 fwd. EF from Amersham
High quality sequence stop: 180.

FEATURES

source

1. .298
/organism="Homo sapiens"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site.1: Not I; Site.2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19w, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following RAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
1.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "
/db_xref="taxon:9606"

BASE COUNT

79 a 50 c 53 g 116 t

ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 298;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;

Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 TTCGAGAA 227
|||||

Oy 25 ttctcagaa 33

RESULT 144
LOCUS AA161311 298 bp mRNA EST 12-MAR-1998
DEFINITION zq38b09.r1 Stratagene hnt neuron (#937233) Homo sapiens cDNA clone
IMAGE:631961.5' similar to gb:M64930 PROTEIN PHOSPHATASE PP2A, 55
KD REGULATORY SUBUNIT, NEURONAL (HUMAN);, mRNA sequence.

ACCESSION AA161311
NID 91735547
VERSION AA161311.1 GI:1735547
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 298)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
On Jan 24, 1995 this sequence version replaced gi:634330.

TITLE JOURNAL
COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1421 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 217.
Location/Qualifiers

FEATURES
source 1..298
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Differentiated, post mitotic hnt neurons. Average insert
size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGACGACGAC 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'."
/db_xref="GDB:5182041"
/db_xref="taxon:9606"
/clone_id="IMAGE:631961"
/clone_1id="Stratagene hnt neuron (#937233)"
/dev_stage="hnt neurons"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 73 a 83 c 75 g 63 t 4 others

ORIGIN
Query Match 100.0%; Score 9; DB 9; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 147 TTCGAGAA 155
|||||

Cp 33 ttctcagaa 25

RESULT 145
LOCUS AA363229 298 bp mRNA EST 21-APR-1997
DEFINITION EST73119 Ovary 1 Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA363229
NID 92015600
VERSION AA363229.1 GI:2015600

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 298)
AUTHORS Adams, M.D., Soares, M.B., Kerlavage, A.R., Fields, C. and Venter, J.C.
Rapid cDNA sequencing (expressed sequence tags) from a
directionally cloned human infant brain cDNA library
Nature Genet. 4, 373-380 (1993)
94004965

JOURNAL
MEDLINE
COMMENT

On Apr 14, 1993 this sequence version replaced gi:692736.
Other ESTs: EST73118 THC167695
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source 1..298
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
/db_xref="ATCC (Inhost):105624"
/db_xref="taxon:9606"
/clone_1id="Ovary 1"
/sex="female"
/dev_stage="adult, 71 yrs"
/dev_stage="adult, 71 yrs"
/dev_stage="adult, 71 yrs"

BASE COUNT 66 a 95 c 73 g 60 t 4 others

ORIGIN
Query Match 100.0%; Score 9; DB 12; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 64 TTCGAGAA 72
|||||

Cp 33 ttctcagaa 25

RESULT 146
LOCUS AA168582 298 bp mRNA EST 12-FEB-1997
DEFINITION m331h06.r1 Stratagene mouse skin (#937213) Mus musculus cDNA clone
IMAGE:608603.5' similar to TR:G1050958 G1050958 CHITROTRIOLIDASE
PRECURSOR.,, mRNA sequence.

ACCESSION AA168582
NID 91746785
VERSION AA168582.1 GI:1746785
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 298)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
On Oct 30, 1996 this sequence version replaced gi:1656875.

TITLE JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:374035
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 278.

FEATURES

source

1. .298
/organism="Mus musculus"
/strain="C57BL/6"
/note="Organ: skin; Vector: Bluescript SK-; Site:1;
ECORI: Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."
/db_xref="taxon:10090"
/clone="IMAGE:608603"
/clone_1lb="Stratagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT

78 a 74 c 69 g 77 t

ORIGIN

Query Match 100.0%; Score 9; DB 9; Length 298;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 116 TTCTCAGAA 124

CP 33 tctccagaa 25

RESULT 147

LOCUS AA977787 299 bp mRNA EST 17-MAR-1999
DEFINITION OG67907.s1 NCI_CGAP_Kid6 Homo sapiens CDNA clone IMAGE:1591452 3',
mRNA sequence.

ACCESSION

AA977787

NID

93155233

VERSION

AA977787.1 GI:3155233

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 299)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE

Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

On Jan 17, 1998 this sequence version replaced gi:1901028.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Stratagene, Inc.

DNA Library Arrayed by: Greg Lennon, Ph.D.

DNA sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www-bio.llnl.gov/dbtrp/image/image.html

Insert length: 489 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES

source

1. .299
/organism="Homo sapiens"

/note="Organ: kidney; Vector: Bluescript SK-; Site:1;
ECORI: Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."
/db_xref="taxon:9606"
/map="846F08"
/clone="IMAGE:1591452"
/clone_1lb="NCI_CGAP_Kid6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT

77 a 57 c 102 g 63 t

ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 2,23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 TTCTCAGAA 225

OY 25 tctccagaa 33

RESULT 148

LOCUS AA639969 299 bp mRNA EST 23-OCT-1997
DEFINITION np12g12.s1 NCI_CGAP_P13 Homo sapiens CDNA clone IMAGE:116166
similar to gb:X56740 RAS-RELATED PROTEIN RAB-11 (HUMAN); mRNA
sequence.

ACCESSION

AA639969

NID

92563748

VERSION

AA639969.1 GI:2563748

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 299)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE

Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

On Dec 30, 1996 this sequence version replaced gi:1529511.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,

M.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Kitzman, Ph.D.

DNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.

DNA sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www-bio.llnl.gov/dbtrp/image/image.html

Trace considered overall poor quality.

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 1.

FEATURES

source

1. .299
/organism="Homo sapiens"
/note="Vector: pAMP10; Site:1: NotI; Site:2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically-determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into pAMP10 by the unc-priming method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by

```

David Krizman."
/db_xref="taxon:9606"
/clone="IMAGE:1116166"
/clone_lib="NCI_CGAP_Pri3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
BASE COUNT      86 a 57 c 70 g      86 t
ORIGIN

Query Match      100.0%; Score 9; DB 16; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 197 TTCGAGAA 205
|||||
25 ttctgagaa 33

RESULT 149
LOCUS      AA336176      299 bp      mRNA      EST      21-APR-1997
DEFINITION EST40702 Epididymus Homo sapiens CDNA 5' end, mRNA sequence.
ACCESSION  AA336176
NID         91988414
VERSION     AA336176.1 GI:1988414
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 299)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulmer,R.A.,
             Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
             White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Val,C.,
             Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
             Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
             Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
             Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
             Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
             Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R.,
             Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
             Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
             Dinko,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
             He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
             Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weisener,P.S., Olsen,H.,
             Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Rubin,S.M.,
             Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
             Fraser,C.M. and Venter,J.C.
TITLE       Initial assessment of human gene diversity and expression patterns
             based upon 83 million nucleotides of cDNA sequence
JOURNAL     Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE     96026280
COMMENT     On Sep 21, 1992 this sequence version replaced g1:276126.
             Other ESTs: THC172355
             Contact: Kerlavage, AR
             Bioinformatics
             The Institute for Genomic Research
             9712 Medical Center Drive, Rockville, MD 20850 USA
             Tel: 3018699056
             Fax: 3018699423
             Email: arkerlav@tigr.org
             For clone availability, additional sequence and expression
             information related to this EST, please check the TIGR Human Gene
             Index (http://www.tigr.org/tdb/hgi/hgi.html)
             Seq primer: M13 Reverse.
FEATURES
source
1..299
    /organism="Homo sapiens"
    /note="organ: epididymis; Vector: pBluescript SK-; Site_1:
    EcoRI; Site_2: XhoI"
    /db_xref="ATCC (inhost):137861"
    /db_xref="taxon:9606"
    /clone_lib="Epididymus"

David Krizman."
/db_xref="taxon:9606"
/clone="IMAGE:1116166"
/clone_lib="NCI_CGAP_Pri3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
BASE COUNT      86 a 57 c 70 g      86 t
ORIGIN

Query Match      100.0%; Score 9; DB 11; Length 299;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 TTCGAGAA 128
|||||
33 ttctgagaa 25

RESULT 150
LOCUS      C38297      300 bp      mRNA      EST      10-SEP-1997
DEFINITION C38297 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
             clone yk490H4 3', mRNA sequence.
ACCESSION  C38297
NID         92374534
VERSION     C38297.1 GI:2374534
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
ORGANISM    Caenorhabditis elegans
             Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
             Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1 (bases 1 to 300)
AUTHORS     Kohara,Y., Motobashiri,T., Tabara,H., Watanabe,H., Sugimoto,A.,
             Sano,M., Miyata,A. and Nishigaki,A.
TITLE       Expression map of the C.elegans genome
JOURNAL     Unpublished (1996)
COMMENT     On Sep 12, 1996 this sequence version replaced g1:1393384.
FEATURES
source
1..300
    /organism="Caenorhabditis elegans"
    /strain="CB1489 him-8(e1489)"
    /note="dev_stage=varied; sex=hermaphrodite male,
    tissue_type=whole animal"
    /db_xref="taxon:6239"
    /clone="yk490H4"
    /clone_lib="Yuj1 Kohara unpublished CDNA"
BASE COUNT      47 a 73 c 75 g 104 t      1 others
ORIGIN

Query Match      100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 18 TTCGAGAA 26
|||||
25 ttctgagaa 33

RESULT 151
LOCUS      C36182      300 bp      mRNA      EST      10-SEP-1997
DEFINITION C36182 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
             clone yk455g1 3', mRNA sequence.
ACCESSION  C36182
NID         92372323
VERSION     C36182.1 GI:2372323
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
ORGANISM    Caenorhabditis elegans
             Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
             Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.

```

REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano,M., Miyata,A. and Nishigaki,A.
 TITLE Expression map of the C.elegans genome
 JOURNAL Unpublished (1996)
 COMMENT On Sep 21, 1992 this sequence version replaced g1:276412.

Contact: Yuji Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 0559-75-0771
 Fax: 0559-75-6240
 Email: ykohara@dbj.nig.ac.jp.

FEATURES
 source 1..300
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /note="dev_stage=varied, sex=hermaphrodite male, tissue_type=whole animal"
 /db_xref="taxon:6239"
 /clone="yk455g1"
 /clone_lib="Yuji Kohara unpublished CDNA"
 /base_count 111 a 49 c 62 g 77 t 1 others

BASE COUNT 111 a 49 c 62 g 77 t 1 others
 ORIGIN

Query Match 100.0% Score 9; DB 15; Length 300;
 Best Local Similarity 100.0% Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 203 TTCGAGAA 211
 ||||||||
 Oy 25 ttctgagaa 33

RESULT 152
 LOCUS C37813 300 bp mRNA EST 10-SEP-1997
 DEFINITION C37813 yuji kohara unpublished cdna Caenorhabditis elegans CDNA
 accession clone yk479a6 3', mRNA sequence.
 C37813
 92374050
 C37813.1 GI:2374050
 EST.
 Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 300)
 Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano,M., Miyata,A. and Nishigaki,A.
 TITLE Expression map of the C.elegans genome
 JOURNAL Unpublished (1996)
 COMMENT On Sep 12, 1996 this sequence version replaced g1:1395428.

Contact: Yuji Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 0559-75-0771
 Fax: 0559-75-6240
 Email: ykohara@dbj.nig.ac.jp.

FEATURES
 source 1..300
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /note="dev_stage=varied, sex=hermaphrodite male, tissue_type=whole animal"
 /db_xref="taxon:6239"
 /clone="yk479a6"
 /clone_lib="Yuji Kohara unpublished CDNA"
 /base_count 47 a 75 c 71 g 105 t 2 others

BASE COUNT 47 a 75 c 71 g 105 t 2 others
 ORIGIN

Query Match 100.0% Score 9; DB 15; Length 300;
 Best Local Similarity 100.0% Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 23 TTCGAGAA 31
 ||||||||
 Oy 25 ttctgagaa 33

RESULT 153
 LOCUS C95112 300 bp mRNA EST 10-JUN-1998
 DEFINITION C95112 pig back fat Sus scrofa CDNA, mRNA sequence.
 accession C95112
 MID 93205085
 VERSION C95112.1 GI:3205085
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Artiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 300)
 AUTHORS Hamanaka,N. and Suzuki,H.
 TITLE Characterization of a porcine fat tissue CDNA library
 JOURNAL Unpublished (1998)
 COMMENT On Sep 12, 1996 this sequence version replaced g1:1405025.

Contact: Noriyuki Hamanaka
 Animal Genome Research Program Team
 STAFF-Institute
 446-1 Ippaiyuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan
 Tel: 81-298-38-2190
 Fax: 81-298-38-2337
 Email: hamanaka@staff.az-jp.
 Location/Qualifiers

FEATURES
 source 1..300
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="pig back fat"
 /tissue_type="back fat"
 /base_count 92 a 62 c 65 g 74 t 7 others
 ORIGIN

Query Match 100.0% Score 9; DB 21; Length 300;
 Best Local Similarity 100.0% Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 166 TTCGAGAA 174
 ||||||||
 Oy 25 ttctgagaa 33

RESULT 154
 LOCUS C33644 300 bp mRNA EST 09-SEP-1997
 DEFINITION C33644 yuji kohara unpublished cdna Caenorhabditis elegans CDNA
 accession clone yk374h2 3', mRNA sequence.
 C33644
 92365440
 C33644.1 GI:2365440
 EST.
 Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 300)
 Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano,M., Miyata,A. and Nishigaki,A.
 TITLE Expression map of the C.elegans genome
 JOURNAL Unpublished (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced g1:692546.

Contact: Yuji Kohara
 Gene Library Lab
 National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.

FEATURES
Source Location/Qualifiers
1. 300

/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/note="dev_stage=varied, sex=Hermaphrodite male,
tissue_type=whole animal"
/db_xref="taxon:6239"
/map="857e09"
/clone="YK374h2"
/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 93 a 55 c 63 g 80 t 9 others
ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 TTCTGAGAA 53
|||||
OY 25 ttctgagaa 33

RESULT 155
LOCUS C30237 300 bp mRNA EST 09-SEP-1997
DEFINITION C30237 Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone YK241a10 3', mRNA sequence.

ACCESSION C30237
NID 92362033
VERSION C30237.1 GI:2362033
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393967.

FEATURES
Source Location/Qualifiers
1. 300
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/note="dev_stage=varied, sex=Hermaphrodite male,
tissue_type=whole animal"
/db_xref="taxon:6239"
/clone="YK241a10"
/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 63 a 70 c 77 g 90 t
ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 253 TTCTGAGAA 261
|||||
OY 25 ttctgagaa 33

RESULT 156
LOCUS C32017 300 bp mRNA EST 09-SEP-1997
DEFINITION C32017 Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone YK314a7 3', mRNA sequence.

ACCESSION C32017
NID 92363813
VERSION C32017.1 GI:2363813
KEYWORDS EST.
SOURCE Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404980.

FEATURES
Source Location/Qualifiers
1. 300
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/note="dev_stage=varied, sex=Hermaphrodite male,
tissue_type=whole animal"
/db_xref="taxon:6239"
/clone="YK314a7"
/clone_lib="Yuji Kohara unpublished cDNA"

BASE COUNT 103 a 47 c 58 g 82 t 10 others
ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 70 TTCTGAGAA 78
|||||
OY 25 ttctgagaa 33

RESULT 157
LOCUS C36735 300 bp mRNA EST 10-SEP-1997
DEFINITION C36735 Yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone YK472c2 3', mRNA sequence.

ACCESSION C36735
NID 92372876
VERSION C36735.1 GI:2372876
KEYWORDS EST.
SOURCE Caenorhabditis elegans.

REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1292408.

Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.

```

FEATURES
  source
    Location/Qualifiers
      1..300
        /organism="Caenorhabditis elegans"
        /strain="CB1489 him-8(e1489)"
        /note="dev_stage=varied, sex=Hermaphrodite male,
        tissue_type=whole animal"
        /db_xref="taxon:6239"
        /clone_lib="Yuj1 Kohara unpublished cDNA"
        /clone_1id="72 c" 72 g 104 t 3 others
BASE COUNT
  49 a 72 c 72 g 104 t
ORIGIN
  Query Match 100.0%; Score 9; DB 15; Length 300;
  Best Local Similarity 100.0%; Pred. No. 2,23e+03;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 18 TTCGAGAA 26
  |||||
  25 ttctgaga 33

RESULT 158
LOCUS C34709 300 bp mRNA EST 10-SEP-1997
DEFINITION C34709 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk415c2 3', mRNA sequence.
ACCESSION C34709
NID 92370850
VERSION C34709.1 GI:2370850
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Apr 14, 1993 this sequence version replaced g1:785592.

Contact: Yuj1 Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.j.
Location/Qualifiers
  1..300
    /organism="Caenorhabditis elegans"
    /strain="CB1489 him-8(e1489)"
    /note="dev_stage=varied, sex=Hermaphrodite male,
    tissue_type=whole animal"
    /db_xref="taxon:6239"
    /map="89602: 15; 15q21.3-15q22.2"
    /clone="Yk415c2"
    /clone_lib="Yuj1 Kohara unpublished cDNA"
BASE COUNT
  91 a 52 c 65 g 92 t
ORIGIN
  Query Match 100.0%; Score 9; DB 15; Length 300;
  Best Local Similarity 100.0%; Pred. No. 2,23e+03;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db .281 TTCGAGAA 289
  |||||
  Cp 33 ttctcaga 25

RESULT 159
LOCUS C31427 300 bp mRNA EST 09-SEP-1997
DEFINITION C31427 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk282f2 3', mRNA sequence.

```

```

ACCESSION C31427
NID 92363223
VERSION C31427.1 GI:2363223
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced g1:1397760.

Contact: Yuj1 Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.j.
Location/Qualifiers
  1..300
    /organism="Caenorhabditis elegans"
    /strain="CB1489 him-8(e1489)"
    /note="dev_stage=varied, sex=Hermaphrodite male,
    tissue_type=whole animal"
    /db_xref="taxon:6239"
    /map="21"
    /clone="Yk282f2"
    /clone_lib="Yuj1 Kohara unpublished cDNA"
BASE COUNT
  51 a 71 c 69 g 105 t 4 others
ORIGIN
  Query Match 100.0%; Score 9; DB 15; Length 300;
  Best Local Similarity 100.0%; Pred. No. 2,23e+03;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 251 TTCGAGAA 259
  |||||
  Cp 33 ttctcaga 25

RESULT 160
LOCUS C54752 300 bp mRNA EST 16-SEP-1997
DEFINITION C54752 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone yk398el 3', mRNA sequence.
ACCESSION C54752
NID 92399353
VERSION C54752.1 GI:2399353
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,
Sano,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced g1:1404644.

Contact: Yuj1 Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.j.
Location/Qualifiers
  1..300
    /organism="Caenorhabditis elegans"

```

```

/strain="CB1489 him-8(e1489)"
/notes="dev-stage=varied, sex=Hermaphrodite male,
tissue_type=whole animal"
/db_xref="taxon:6239"
/map="MMU16C3-C4 region"
/clone="YK39861"
/clone_lib="Yuji Kohara unpublished cDNA"
BASE COUNT      118 a      38 c      64 g      78 t      2 others
ORIGIN

Query Match      100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      166 TTCTCAGAA 69
        |||||
        33 ttctcagaa 25

RESULT 161
LOCUS   C56383      300 bp      mRNA      EST      16-SEP-1997
DEFINITION C56383 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone YK261a12 3', mRNA sequence.
ACCESSION C56383
VERSION   92400984
KEYWORDS  C56383.1 GI:2400984
SOURCE    EST.
ORGANISM  Caenorhabditis elegans.
REFERENCE 1 (bases 1 to 300)
AUTHORS   Kohara, Y., Motokashi, T., Tabara, H., Watanabe, H., Sugimoto, A.,
Sano, M., Miyata, A. and Nishigaki, A.
TITLE      Expression map of the C.elegans genome
JOURNAL    Unpublished (1996)
COMMENT    On Apr 14, 1993 this sequence version replaced gi:692692.

Contact: Yuji Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.
Location/Qualifiers
1. 300
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/notes="dev-stage=varied, sex=Hermaphrodite male,
tissue_type=whole animal"
/db_xref="taxon:6239"
/clone="YK261a12"
/clone_lib="Yuji Kohara unpublished cDNA"
BASE COUNT      110 a      49 c      63 g      76 t      2 others
ORIGIN

Query Match      100.0%; Score 9; DB 15; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      166 TTCTCAGAA 174
        |||||
        33 ttctcagaa 25

RESULT 162
LOCUS   AA382476      300 bp      mRNA      EST      21-APR-1997
DEFINITION EST95712 Testis I Homo sapiens cDNA 5' end similar to
Proto-oncogene BMI-1, mRNA sequence.
ACCESSION AA382476
VERSION   92034815
KEYWORDS  AA382476.1 GI:2034815
SOURCE    EST.
ORGANISM  Caenorhabditis elegans.

```

```

KEYWORDS EST.
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS   Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 300)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, M.C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.,
Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palanges, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruden, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
Fraser, C.M. and Venter, J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
JOURNAL    Nature
MEDLINE    96026280
COMMENT    On Sep 1, 1995 this sequence version replaced.
Other ESTs: THC175599
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavage@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.cigr.org/tadb/ngl/ngl.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 300
/organism="Homo sapiens"
/notes="Organ: testis; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI"
/db_xref="AFCC (Inhost):186783"
/db_xref="taxon:9606"
/map="10"
/clone_lib="Testis I"
/sex="male"
/dev_stage="adult"
BASE COUNT      113 a      38 c      71 g      75 t      3 others
ORIGIN

Query Match      100.0%; Score 9; DB 12; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      207 TTCTCAGAA 215
        |||||
        33 ttctcagaa 25

RESULT 163
LOCUS   C54178      300 bp      mRNA      EST      11-SEP-1997
DEFINITION C54178 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA
clone YK353g7 3', mRNA sequence.
ACCESSION C54178
VERSION   92391935
KEYWORDS  C54178.1 GI:2391935
SOURCE    EST.
ORGANISM  Caenorhabditis elegans.

```

Eukaryota; Metazoa; Nematozoa; Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis. 1 (bases 1 to 300)
 Kohara, Y., Motohashi, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano, M., Miyata, A. and Nishigaki, A.
 Expression map of the C. elegans genome
 Unpublished (1996)
 On Apr 14, 1993 this sequence version replaced gl:716864.

Contact: Yuji Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 0559-75-0771
 Fax: 0559-75-6240
 Email: ykohara@dbj.nig.ac.jp.
 Location/Qualifiers

FEATURES

1..300
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /note="dev_stage=varied, sex=Hermaphrodite male, tissue_type=whole animal"
 /db_xref="taxon:6239"
 /clone="YK53367"
 /clone_1lb="Yuji Kohara unpublished cDNA"
 BASE COUNT 92 a 70 c 44 g 94 t
 ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 173 TTCTCAGAA 181
 |||||
 Cp 33 ttctcagaa 25

RESULT 164
 LOCUS AA516287 300 bp mRNA EST 19-AUG-1997
 DEFINITION U68104.s1 NCI_CGAP_C03 Homo sapiens cDNA clone IMAGE:925111 3', mRNA sequence.
 ACCESSION AA516287
 NID 92255887
 VERSION AA516287.1 GI:2255887
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 300)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 On Apr 5, 1995 this sequence version replaced gl:760509.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arraying: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbirp/image/image.html

Insert Length: 522 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 203.
 Location/Qualifiers
 1..300

/organism="Homo sapiens"
 /note="vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization.
 /db_xref="taxon:9606"
 /clone="IMAGE:925111"
 /clone_1lb="NCI_CGAP_C03"
 /sex="pooled"
 /tissue_type="colon"
 /lab_host="DH10B"
 BASE COUNT 105 a 53 c 69 g 73 t
 ORIGIN

Query Match 100.0%; Score 9; DB 14; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 255 TTCTCAGAA 263
 |||||
 Cp 33 ttctcagaa 25

RESULT 165
 LOCUS C53705 300 bp mRNA EST 11-SEP-1997
 DEFINITION C53705 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA clone YK323d6 3', mRNA sequence.
 ACCESSION C53705
 NID 92391462
 VERSION C53705.1 GI:2391462
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematozoa; Secernentea; Rhabditia; Rhabditidae; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 300)
 Kohara, Y., Motohashi, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano, M., Miyata, A. and Nishigaki, A.
 Expression map of the C. elegans genome
 Unpublished (1996)
 On Apr 14, 1993 this sequence version replaced gl:693591.

Contact: Yuji Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 0559-75-0771
 Fax: 0559-75-6240
 Email: ykohara@dbj.nig.ac.jp.
 Location/Qualifiers

FEATURES

1..300
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /note="dev_stage=varied, sex=Hermaphrodite male, tissue_type=whole animal"
 /db_xref="taxon:6239"
 /map="17; 4"
 /clone="YK323d6"
 /clone_1lb="Yuji Kohara unpublished cDNA"
 BASE COUNT 119 a 38 c 64 g 79 t
 ORIGIN

Query Match 100.0%; Score 9; DB 15; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2,23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 65 TTCTCAGAA 73
 |||||
 Cp 33 ttctcagaa 25

RESULT 166
 LOCUS C53842 300 bp mRNA EST 11-SEP-1997
 DEFINITION C53842 Yuj1 Kohara unpublished cDNA *Caenorhabditis elegans* CDNA
 clone YK31745 3, mRNA sequence.
 ACCESSION C53842
 MID g2391599
 VERSION C53842.1 GI:2391599
 KEYWORDS EST.
 SOURCE *Caenorhabditis elegans*.
 ORGANISM *Caenorhabditis elegans*.
 Eukaryota; Metazoa; Chordata; Secernentea; Rhabdittia; Rhabdittida;
 Rhabdittina; Rhabdittolidae; Rhabdittidae; Peloderinae; *Caenorhabditis*.
 REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara, Y., Mochizuki, T., Tabara, H., Matsuura, H., Sugimoto, A.,
 Sano, M., Miyata, A. and Nishigaki, A.
 TITLE Expression map of the *C. elegans* genome
 JOURNAL Unpublished (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced g1:716436.

CONTACT: Yuj1 Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 055-75-0771
 Fax: 055-75-6240
 Email: ykohara@db.nig.ac.jp.
 Location/Qualifiers
 1. 300
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /note="dev-stage=varied; sex=Hermaphrodite male;
 tissue_type=whole animal"
 /db_xref="taxon:6239"
 /map="871E01: 5: 5q21.3"
 /clone="YK31745"
 /clone_1bp="Yuj1 Kohara unpublished CDNA"
 /clone_1bp="32 g 103 t
 36 c 32 g 103 t

BASE COUNT 129 a 36 c 32 g 103 t
 ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 48 TTCACGAA 56
 |||||
 Cp 33 ttccagaa 25

RESULT 167
 LOCUS C94621 300 bp mRNA EST 10-JUN-1998
 DEFINITION C94621 Pig back fat *Sus scrofa* CDNA, mRNA sequence.
 ACCESSION C94621
 MID g3204594
 VERSION C94621.1 GI:3204594
 KEYWORDS EST.
 SOURCE *Sus scrofa*.
 ORGANISM *Sus scrofa*.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Artiodactyla; Suidae; Suidae; Suidae; Suidae.
 REFERENCE 1 (bases 1 to 300)
 AUTHORS Hamada, N. and Suzuki, H.
 TITLE Characterization of a porcine fat tissue cDNA library
 JOURNAL Unpublished (1998)
 COMMENT On Jan 9, 1998 this sequence version replaced g1:936835.

CONTACT: Noriyuki Hamada
 Animal Genome Research Program Team
 STAFF-Institute
 446-1 Ipaizuka, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan
 Tel: 81-298-38-2190
 Fax: 81-298-38-2337

FEATURES
 source
 Location/Qualifiers
 1. 300
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /map="21"
 /clone_1bp="Pig back fat"
 /tissue_type="back fat"

BASE COUNT 100 a 44 c 68 g 88 t
 ORIGIN

Query Match 100.0%; Score 9; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 275 TTCACGAA 283
 |||||
 Cp 33 ttccagaa 25

RESULT 168
 LOCUS AA385118 301 bp mRNA EST 21-APR-1997
 DEFINITION EST98723 Thyroid Homo sapiens CDNA 5' end similar to
 NM001001000 oxidoreductase B22 subunit, mRNA sequence.
 ACCESSION AA385118
 MID g2037437
 VERSION AA385118.1 GI:2037437
 KEYWORDS EST.
 SOURCE human.
 ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 301)
 AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
 Bult, C.J., Lee, N.H., Kirkness, E.F., Weissbrock, K.G., Gocayne, J.D.,
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, M.L.C.,
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
 Fitzgerald, L.M., Fritch, W.M., Fritch, J.L., Geoghegan, N.S.,
 Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., J.R.,
 Kelley, J.M., Kelley, J.C., Liu, L.-I., Marra, S.M., Merrick, J.M.,
 Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Peligro, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Wiley, J.,
 Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dimke, D., Feng, D.F., Ferris, A., Fischer, C., Hastings, G.A.,
 He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
 Kozak, D.L., Kunsch, C., Hung, J., Li, H., Melissner, P.S., Olsen, R.,
 Raymond, L., Welty, F., Ming, J., Xu, C., Yu, G.L., Ruben, S.M.,
 Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
 Fraser, C.M. and Venter, J.C.
 TITLE Initial assessment of human gene diversity and expression patterns
 JOURNAL based upon 83 million nucleotides of cDNA sequence
 MEDLINE Nature 377 (6547 Suppl), 3-174 (1995)
 COMMENT 96026280
 On Apr 14, 1993 this sequence version replaced g1:765863.
 Other ESTs: THG128658
 CONTACT: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavage@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. 301
 /organism="Homo sapiens"
 /note="Organ: thyroid gland; Vector: pBluescript SK-"
 /site_1: EcorI; Site_2: XhoI
 /db_xref="ATCC (lnost):189283"

```

/ab_xref="taxon:9606"
/map="9"
/clone_lib="Thyroid"
/dev_stage="adult"
BASE COUNT      80 a      71 c      82 g      66 t      2 others
ORIGIN

Query Match      100.0%; Score 9; DB 12; Length 301;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      287 TTCTGAGAA 295
OY      25 ttctgagaa 33

RESULT 169
LOCUS      AA333727      301 bp      mRNA      EST      21-APR-1997
DEFINITION EST37844 Embryo, 9 week Homo sapiens CDNA 5' end, mRNA sequence.
ACCESSION  AA333727
NID         91985971
VERSION     AA333727.1 GI:1985971
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 301)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirsnes,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man'hai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkley,P.S., Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.F., Ferrle,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meltsner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Ming,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl.), 3-174 (1995)
JOURNAL     96026280
MEDLINE
COMMENT     On Apr 14, 1993 this sequence version replaced gi:693230.
            Other ESTs: THC184626
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018698423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tcdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES
Source
1..301
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
/db_xref="ATCC (inhost):135392"
/db_xref="taxon:9606"
/map="1"
/clone_lib="Embryo, 9 week"
/dev_stage="embryo, 9 wks"
BASE COUNT      91 a      73 c      73 g      61 t      3 others

```

```

ORIGIN

Query Match      100.0%; Score 9; DB 11; Length 301;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      263 TTCTGAGAA 271
OY      25 ttctgagaa 33

RESULT 170
LOCUS      A1019762      301 bp      mRNA      EST      16-JUN-1998
DEFINITION ua90g04.r1 Soares mouse mammary gland NMIMG Mus musculus cDNA clone
IMAGE:1364790 5', mRNA sequence.
ACCESSION  A1019762
NID         93234098
VERSION     A1019762.1 GI:3234098
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 301)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schallenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMT Mouse EST Project
            Unpublished (1996)
JOURNAL     On Jan 17, 1998 this sequence version replaced gi:2044829.
COMMENT

TITLE
JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:898010
Putative full length read
vector to vector length is 302
Seq primer: -28m3 rev2 ET from Amersham.
FEATURES
Source
1..301
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTTCACATCTGAGAGTGAGGAGCGCCCGAATGTTTCTTTTCTTTTCTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
/db_xref="taxon:10090"
/map="Chromosome 11"
/clone_lib="Soares mouse mammary gland NMIMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT      84 a      48 c      58 g      111 t
Query Match      100.0%; Score 9; DB 21; Length 301;

```

Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 150 TTCTCAGAA 158
|||||
Cp 33 ttctcagaa 25

RESULT 171

LOCUS AA381889 302 bp mRNA EST 21-APR-1997
DEFINITION EST95202 Activated T-cells 1 Homo sapiens cDNA 5' end, mRNA

ACCESSION AA381889

NID 92034373

VERSION AA381889.1 GI:2034373

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE Adams,M.D., Keriavage,A.R., Fleischmann,R.D., Fulton,R.A.,

1 (bases 1 to 302)

Bult,C.J., Lee,N.H., Kirkness,E.F., Weissstock,K.G., Gocayne,J.D.,

White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wei,C.,

Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,

Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.,

Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,

Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,

Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,

Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,

Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,

Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,

Dinke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,

He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,

Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Melsner,P.S., Olsen,H.,

Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,

Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,

Fraser,C.M., and Venter,J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026580

On Sep 12, 1996 this sequence version replaced gi:1407216.

Other-ESTs: THC168705

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..302

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI"

/db_xref="ATCC (inhost):186435"

/db_xref="taxon:9606"

/clone_lib="Activated T-cells 1"

/cell_type="T-Lymphocyte"

/dev_stage="adult"

/dev_stage="adult"

BASE COUNT 74 a 76 c 72 g 76 t 4 others

ORIGIN

Query Match 100.0%; Score 9; DB 12; Length 302;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 128 TTCTCAGAA 136

Oy 25 ttctcagaa 33
|||||

RESULT 172

LOCUS AA626881 302 bp mRNA EST 15-OCT-1997
DEFINITION z89h12.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745223

3', mRNA sequence.

ACCESSION AA626881

NID 92539268

VERSION AA626881.1 GI:2539268

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,

1 (bases 1 to 302)

Krizek,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,

Martin,J., Moore,B., Scheinberg,K., Steptoe,M., Tan,F.,

Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

Unpublished (1997)

On Sep 12, 1996 this sequence version replaced gi:1288460.

COMMENT

CONTACT: Wilson R

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40m13 fwd. EP from Amersham

High quality sequence stop: 299.

Location/Qualifiers

1..302

/organism="Homo sapiens"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from Clontech

Laboratories, Inc., and primed with a Not I - oligo(dT)

primer [5']

TGTTACCAATCTGACATGCGAGCGCCGCCCAATTTTTTTTTTTT 3']

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization to Cot5, and was

constructed by Bento Soares and M. Fatima Bonaldo.

/db_xref="GDB:5933093"

/db_xref="taxon:9606"

/clone_image:745223"

/clone_lib="Soares_testis_NHT"

/sex="male"

/lab_host="DH10B"

BASE COUNT 87 a 45 c 48 g 122 t

ORIGIN

Query Match 100.0%; Score 9; DB 16; Length 302;

Best Local Similarity 100.0%; Pred. No. 2.23e+03;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 40 TTCTCAGAA 48

|||||

Cp 33 ttctcagaa 25

RESULT 173

LOCUS AA712997 302 bp mRNA EST 24-DEC-1997

DEFINITION 32557 Lambda-PRL2 Arabidopsis thaliana cDNA clone 220L22T7, mRNA

sequence.

ACCESSION AA712997

NID 92722914

VERSION AA712997.1 GI:2722914
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 Rosidae; Caprales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 302)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
 Rezel,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1397889.
 COMMENT

TITLE
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1397889.
 COMMENT

FEATURES
 source
 1. .302
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /note="Vector: lambda zip-lox; Site.1: Sal; Site.2: Not;
 lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark-rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda zip-lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA."
 /db_xref="taxon:3702"
 /map="18q22.3-qter"
 /clone="220L22T7"
 /clone_1lb="lambda-PRL2"
 BASE COUNT 86 a 53 c 73 g 78 t 12 others
 ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 52 TTCGACAA 60
 ||||||||
 CP 33 ttctcagaa 25

RESULT 174
 LOCUS T14558 304 bp mRNA EST 27-FEB-1995
 DEFINITION SMESS10219 Schistosoma mansoni, adult worm, Gloria Franco
 Schistosoma mansoni cDNA clone SMPA159 3' end, mRNA sequence.
 T14558
 ACCESSION 9514725
 NID 9514725
 VERSION T14558.1 GI:514725
 KEYWORDS EST.
 ORGANISM Schistosoma mansoni.
 Eukaryote: Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigidae; Schistosomatidae; Schistosomatidae; Schistosoma.
 1 (bases 1 to 304)
 Franco,G.R., Adams,M.D., Soares,M.B., Simpson,A.J.G., Venter,J.C.
 and Pena,S.D.J.
 Identification of new Schistosoma mansoni genes by the EST strategy
 using a directional cDNA library

JOURNAL Gene 152, 141-147 (1995)
 MEDLINE 95137379
 COMMENT

CONTACT: Franco G.R. and Pena S.D.J.
 Laboratorio de Genetica-Bioquimica, Departamento de Biologia
 Imunologia
 Instituto de Ciencias Biologicas, Universidade Federal de Minas
 Gerais
 Avenida Antonio Carlos 6627, Belo Horizonte, MG, Brazil, 31270-010
 Tel: (5531)4415611
 Fax: (5531)4415409
 Email: gfranco@monoc.icb.ufmg.br
 Seq primer: M13 Forward
 Location/Qualifiers
 1. .304
 /organism="Schistosoma mansoni"
 /strain="NMRI"
 /note="Vector: BA vector; Site.1: NotI; Site.2: HindIII;
 Total cellular RNA from male and female adult worms was
 extracted according to a modification (Puissant, C. and
 Gaudelme, L. M. Biofeedback 8, 148-149, 1990) of the
 Guanidine Thiocyanate procedure (Chomczynski, P. and
 Sacchi, N. Anal. Biochem. 162, 156-159, 1987). Poly (A)+
 RNA was purified by oligo dt column and cDNA was
 synthesized as described previously (Adams, M. D. et al.
 Nature Genet. 4, 373-389, 1993). cDNA was ligated to a
 two fold molar excess of a NotI/HindIII digested plasmid
 DNA (lambda BA vector, a phagemid derived from pEMBL,
 Adams, M. D. et al. Nature Genet. 4, 373-389, 1993) and
 electroporated into E. coli strain DH10B (BRL). The
 library was amplified and further selected for clones
 containing long inserts (>500 bp) by purification of the
 plasmid DNA from a fragment of a 1% low-melting-point
 agarose gel, containing the smear of the library and
 electroporation into DH10B cells.
 /db_xref="taxon:6183"
 /map="X"
 /clone="SMPA159"
 /clone_1lb="Schistosoma mansoni, adult worm, Gloria
 Franco"
 /lab_host="DH10B, JM109"
 BASE COUNT 103 a 56 c 55 g 90 t
 ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 304;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 116 TTCGACAA 124
 ||||||||
 CY 25 ttctcagaa 33

RESULT 175
 LOCUS AA992859 304 bp mRNA EST 21-DEC-1998
 DEFINITION ct48h11.s1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:1620069
 3', mRNA sequence.
 AA992859
 ACCESSION 95178404
 NID 95178404
 VERSION AA992859.1 GI:3179404
 KEYWORDS EST.
 ORGANISM human.
 Homo sapiens
 Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 304)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Apr 14, 1993 this sequence version replaced gi:503472.
 Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bernaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.1lnl.gov/bdrp/image/image.html

Insert length: 657 Std Error: 0.00
 Seq primer: -40ml3 fwd. RT from Amersham
 High quality sequence stop: 292.

FEATURES
 source
 Location/Qualifiers

1. 304
 /organism="Homo sapiens"
 /note="Vector: pT73-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5']
 TGTACCAATCTGAGTGGAGCGCCGCCCAATTTTCTTTT 3'1.
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cots, and was constructed by Bento Soares and M. Fatima Bernaldo.
 /db_xref="taxon:9606"
 /clone="IMAGE:1620069"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /lab="host"

BASE COUNT 114 a 51 c 51 g 88 t
 ORIGIN

Query Match 100.0% Score 9; DB 21; Length 304;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 208 TTCTGAGAA 216
 |||||
 Cp 33 ttctcgaga 25

RESULT 176
 LOCUS AA313179 305 bp mRNA EST 19-APR-1997
 DEFINITION EST18492 Liver, subtracted (abundant clones) I Homo sapiens CDNA 5' end similar to similar to albumin, mRNA sequence.
 ACCESSION AA313179
 NID g1965528
 VERSION AA313179.1 GI:1965528
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Adams, M.D., Keriavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-vel, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodde, R.A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palanges, R.F., McDonald, L.A., Nguyen, D.T., Peligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y., Beharlik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dinke, D., Feng, D.-F., Fertile, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Melsen, P.S., Olsen, H., Raymond, L.V., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,

TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Fraser, C.M. and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl.), 3-174 (1995)
 96026280
 On Apr 14, 1993 this sequence version replaced g1:692643.

Contact: Keriavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: ar@rivergate.org
 For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source
 Location/Qualifiers

1. 305
 /organism="Homo sapiens"
 /note="Organ: liver; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
 /db_xref="ATCC (lnhost):117113"
 /db_xref="taxon:9606"
 /map="22; 804C11; 5' 5q23.3-5q31.3"
 /clone_lib="Liver, subtracted (abundant clones) I"
 /dev_stage="adult"
 /dev="stage"

BASE COUNT 109 a 57 c 68 g 71 t
 ORIGIN

Query Match 100.0% Score 9; DB 11; Length 305;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 TTCTGAGAA 23
 |||||
 Cp 25 ttctcgaga 33

RESULT 177
 LOCUS AA760552 305 bp mRNA EST 23-JAN-1998
 DEFINITION vv78e11.r1 Stragene mouse skin (#93713) Mus musculus cDNA clone IMAGE:1228556 5' similar to gb:568836 S-ADENOSYLMETHIONINE SYNTHETASE GAMMA FORM (HUMAN);, mRNA sequence.
 ACCESSION AA760552
 NID g2808346
 VERSION AA760552.1 GI:2808346
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (Bases 1 to 305)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 On Sep 12, 1996 this sequence version replaced g1:1405065.

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1810
 Fax: 314 286 1800
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.1lnl.gov) for further information.
 MGI:654148

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 296.
Location/Qualifiers

1. 305

/organism="Mus musculus"

/strain="C57BL/6"

/note="Organ: Skin; Vector: pBluescript SK-; Site: 1:
ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

/db_xref="taxon:10090"

/clone="IMAGE:1228556"

/clone_1lb="Stratagene mouse skin (#937313)"

/tissue_type="whole skin"

/sex="females"

/dev_stage="11 weeks old"

/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 76 a 57 c 77 g 95 t

ORIGIN

Query Match 100.0%; Score 9; DB 18; Length 305;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 TTCTCAGAA 103
|||||

Cp 33 ttctcagaa 25

RESULT 178
LOCUS AA313174 306 bp mRNA EST 19-APR-1997
DEFINITION EST18461 Liver, subtracted (abundant clones) I Homo sapiens CDNA 5'

end similar to albumin, mRNA sequence.

AA313174
91965502

VERSION AA313174.1 GI:1965502

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fulton, R.A.,
Balt, C.J., Lee, N.H., Kinkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, M.L.C.,
Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.,
Gjodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,
Kellay, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,
Moreno-Palmer, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,
Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,
Raymond, L., Welty, F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,
Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,
Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL MEDLINE
COMMENT 96026280

On Apr 14, 1993 this sequence version replaced g1:692610.
Other ESTs: THC172063

Contact: Kerlavage, AR
Bioinformatics

*The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423

Email: arterlavetlgr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgl/hgl.html>)
Seq primer: M13 Reverse.
Location/Qualifiers

1. 306

/organism="Homo sapiens"

/note="Organ: Liver; Vector: pBluescript SK-; Site: 1:
ECORI; Site: 2: XhoI"

/db_xref="ATCC (Inhost):117087"

/db_xref="taxon:9606"

/map="972H07: 12: 12q15-q21.2"

/clone_1lb="Liver subtracted (abundant clones) I"

/dev_stage="adult"

BASE COUNT 99 a 63 c 74 g 69 t 1 others

ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 102 TTCTCAGAA 110
|||||

Cp 25 ttctcagaa 33

RESULT 179
LOCUS AA881995 306 bp mRNA EST 26-MAR-1998
DEFINITION vx33910.r1 Stratagene mouse lung 937302 Mus musculus CDNA clone

IMAGE:1277058 5', mRNA sequence.

AA881995
92980085

VERSION AA881995.1 GI:2990085

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 306)

Warrar, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubouque, T.,
Geisel, S., Kuwaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

The WashU-HMT Mouse EST Project
Unpublished (1996)

On Jan 9, 1998 this sequence version replaced g1:785865.

TITLE

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.lbnl.gov) for further information.
MGI:668858

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 297.
Location/Qualifiers

1. 306

/organism="Mus musculus"

/strain="C57BL/6 x CBA"

/note="Organ: Lung; Vector: pBluescript SK-; Site: 1:
ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. 6-8 month old female lung and 1.5 year old male
lung were source of mRNA. Average insert size: 1.5 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

/db_xref="taxon:10090"

/clone="IMAGE:1277058"

```

1:
/clone_lib="Stratagene mouse lung 937302"
/sex="female"
/tissue_type="lung"
/dev_strage="5-8 month old"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT      86 a      72 c
ORIGIN           71 c      77 g

```

Query Match	100.0%;	Score 9;	DB 20;	Length 306;
Best Local Similarity	100.0%;	Pred. No. 2.23e+03;		
Matches	9;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Db	285	TTCTGAGAA	293
QY	25	ttctgagaa	33

LOCUS	DEFINITION	EST	12-MAR-1995
F11767	306 bp mRNA		
HSC3IC11	normalized infant brain cDNA		
C-3IC11	Homo sapiens cDNA clone		
	c-3IC11, mRNA sequence.		

JOURNAL MEDLINE COMMENT	C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995) 95277534

```

/organism="Homo sapiens"
/name="Organ: brain; Vector: lafmid BA; site_1: HindIII;
site_2: NotI; sex=female; dev.stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B Soares, Psychiatry
Dept, Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
$db_xref="taxon:9606"

```

```
Query Match      100.0%; Score 9; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
```

Matches	9;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	273	TTTCACAGAA	281						
Cp	33	tttcacagaa	25						

RESULT	181			
LOCUS	L25482	308 bp	mRNA	EST 23-SEP-1994
DEFINITION	HDD21S151 Clontech HL 1055a Homo sapiens CDNA clone 21S051, mRNA sequence.			
ACCESSION	U55493			

JOURNAL Genomics 23, 75-84 (1994)
MEDLINE 95130115
COMMENT

```
/organism="Homo sapiens"  
/note="Whole brain, from 26 week old male fetus"  
/db_xref="taxon:9606"
```

Query Match	100.0%;	Score 9;	DB 8;	Length 306;
Best Local Similarity	100.0%;	Pred. No. 2.23e+03;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Db	211	TTCTCAGAA	219
Cp	33	ttctcagaa	25

RESULT	182			
LOCUS	F07471	308 bp	MRNA	EST 20-FEB-1995
DEFINITION	HSC27A041 normalized infant brain cDNA Homo sapiens cDNA clone			
	C-27A04, mRNA sequence.			

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-27a04
Insert Length: 737 Std Error: 0.00
Seq primer: (-21)M13.universal
High quality sequence stop: 349.
Location/Qualifiers

FEATURES

source

1..308
/organism="Homo sapiens"
/note="Organ: Brain; Vector: latmid BA; Site_1: HindIII;
Site_2: NotI; sex=Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
latmid BA vector. Clone library from B.Soaes, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
/db_xref="taxon:9606"
/clone="c-27a04"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
68 c 79 g 80 t 5 others

BASE COUNT

76 a

ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 266 TTCTCAGAA 274

Cp 33 ttctcagaa 25

RESULT 183

LOCUS T90079 308 bp mRNA EST 20-MAR-1995
DEFINITION Y339D06.s1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone
IMAGE:110579.3' similar to gb:M31303_rnal STATHMIN (HUMAN);, mRNA
sequence.

ACCESSION

T90079

NID

9718592

VERSION

T90079.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 308)

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Womlmann,P. and
Wilson,R.

TITLE The Mashu-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 693
High quality sequence stops: 282 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 693 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 282.
Location/Qualifiers

FEATURES

source

1..308
/organism="Homo sapiens"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
(5' AACGTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'),
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
/db_xref="GDB:465196"
/db_xref="taxon:9606"
/clone="IMAGE:110579"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
77 c 66 g 114 t 1 others

BASE COUNT

50 a

ORIGIN

Query Match 100.0%; Score 9; DB 19; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 TTCTCAGAA 34

Cp 33 ttctcagaa 25

RESULT 184

LOCUS AA419174 309 bp mRNA EST 16-OCT-1997
DEFINITION ZV34G05.r1 Soares ovary tumor NBHOT Homo sapiens CDNA clone
IMAGE:755576.5', mRNA sequence.

ACCESSION

AA419174

NID

92078921

VERSION

AA419174.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 309)

AUTHORS

Hillier,L., Allen,M., Bowles,L., Dubuque,T., Getsel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE Mashu-Merck EST Project 1997

JOURNAL Unpublished (1997)

COMMENT On Sep 12, 1996 this sequence version replaced gi:1392948.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 255.
Location/Qualifiers

FEATURES

source

1..309
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTCACATCTGAGTGGAGCGCGCGGCTTTTCTTTTCTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI

Db	228	TTCTGAGAA	236
QY	25	ttctgagaa	33

RESULT 187
LOCUS AA502321 310 bp mRNA EST 19-AUG-1997
DEFINITION ne26606.s1 NCI-CGAP_C03 Homo sapiens cDNA clone IMAGE:598522 3'
similar to WP:F45E12.4 CE02740 ;, mRNA sequence.
ACCESSION AA502321
NID g2237288
VERSION AA502321.1 GI:2237288
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 310)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407174.

BASE COUNT	68	a	82	c	83	g	77	t		
ORIGIN										
Query Match	100.0%; Score 9; DB 14; Length 310;									
Best Local Similarity	100.0%; Pred. No. 2.23e+03;									
Matches	9; Conservative									
	0; Mismatches 0; Indels 0; Gaps 0;									
Db	99	TTCTCAGAA	107							
Cp	33	ttctcagaa	25							
RESULT	188									
LOCUS	AA146289	311	bp	mrna	EST	16-FEB-1997				
DEFINITION	nt00507.r1 Soares mouse 3nbms Mus musculus cDNA clone IMAGE:596700									
	5', mRNA sequence.									
ACCESSION	AA146289									

NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA1715930
GI:1715930
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 311)
Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gastei,S., Klicada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Thaisang,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The MashU-HMI Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced gi:1407419.

ORIGIN	BASE COUNT	110 a	48 c	52 g	101 t
Query Match			100.0%	Score 9;	DB 9;
Best Local Similarity	100.0%		Pred. No. 2.22e+03;	Length 311;	
Matches	9;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Db	260	TTCTCAGAA	268		
Cp	33	ttctcaga	25		
RESULT	189				
LOCUS	T32011	312 bp	mRNA	EST	06-SEP-1995
DEFINITION	EST42478 Human Uterus Homo sapiens cDNA 5' end similar to similar				
	to gastrin-binding protein, mRNA sequence.				
ACCESSION	T32011				
NID	9644109				
VERSION	T32011.1	GI:614109			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				

REFERENCE
AUTHORS

Eutheria: Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 312)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,
Balt, C.J., Lee, N., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D.,
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M.-W.,
Clayton, R.A., Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S.M.,
Glodde, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S.,
Kelleys, J.M., Klimke, K.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M.,
Merrill, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T.,
Pellegrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L.,
Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R.,
Weidman, J.F., Li, Y., Bedarrik, D.P., Cao, L., Cepeda, M.A.,
Coleman, T.A., Collins, E.-J., Dime, D., Feng, P., Ferris, A.,
Fischer, C., Hastings, G.A., He, W.-W., Hu, J.-S., Greene, J.M.,
Gruber, J., Hudson, P., Kim, A., Kozak, D.L., Kunsch, C., Li, H., Li, H.,
Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.-F., Wang, J., Xu, C.,
Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A.,
Hasselstine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 83 Million Basepairs of cDNA Sequence
Nature 377, 3-174 (1995)

JOURNAL
MEDLINE
COMMENT

Other ESTs: THC9254
Contact: Venter, J.C.
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423

Email: tdbinfo@tdb.tigr.org
For clone availability, additional sequence and expression
information related to this EST, please contact the TIGR Database
(tdbinfo@tdb.tigr.org)
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..312
/organism="Homo sapiens"
/note="Organ: uterus"
/db_xref="AFCC (Inhost):102354"
/db_xref="taxon:9606"
/clone_id="Human Uterus"
75 c 71 g 86 t 2 others

BASE COUNT
ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 236 TTCACGAA 244
Cp 33 ttctcagaa 25

RESULT 190
LOCUS AA724581 312 bp mRNA EST 31-DEC-1998
DEFINITION A101b11.s1 Soares_NFL.T.GBC.S1 Homo sapiens cDNA clone
IMAGE:1327293 3', mRNA sequence.

ACCESSION
NID 92742288
VERSION AA724581.1 GI:2742288
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 312)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

JOURNAL
COMMENT

Unpublished (1997)
On Jun 18, 1996 this sequence version replaced gi:1366624.

Contact: Robert Strausberg, Ph.D.

FEATURES

Location/Qualifiers
1..312
/organism="Homo sapiens"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung BHL19W, testis NRT, and B-cell
NCI-CGAP GC81) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/map="11"
/clone="IMAGE:1327293"
/clone_id="Soares_NFL.T.GBC.S1"
/lab_host="DH10B"

BASE COUNT 69 a 65 c 44 g 134 t
ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 143 TTCACGAA 151
Cp 33 ttctcagaa 25

RESULT 191
LOCUS AA713645 313 bp mRNA EST 18-FEB-1998
DEFINITION nv80h11.s1 NCI-CGAP_Bx4 Homo sapiens cDNA clone IMAGE:1236165, mRNA
sequence.

ACCESSION
NID 92725919
VERSION AA713645.1 GI:2725919
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 313)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1404722.

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Ian Kirsch, M.D., Kristina A. Cole, M.D.,
Ph.D. student, Rodrigo F. Chuagui, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 441 Std Error: 0.00

Seq primer: -40m13 fwd. EV from Amersham.

BASE COUNT 86 a 59 c 64 g 106 t
ORIGIN
Query Match 100.0%; Score 9; DB 14; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 194 TTCGACAA 202
|||||
Cp 33 ttctcagaa 25

RESULT 194
LOCUS AA252619 316 bp mRNA EST 13-AUG-1997
DEFINITION zs14d06.s1 NCI_CGAP_GCB1 Homo sapiens CDNA clone IMAGE:685163 3',
mRNA sequence.
ACCESSION AA252619
NID 91887602
VERSION AA252619.1 GI:1887602
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 316)
AUTHORS NCI_CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced gi:693458.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Insert Length: 1076 Std Error: 0.00
Seq primer: -41ml3 fwd: ET from Amersham
High quality sequence stop: 315.
Location/Qualifiers
1. 316
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, 19D-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
15'-TGTTACCATCTGAGTGGAGCGGCGCTCATTTTCTTTTCTTTT-
3'. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/map="8"
/clone="IMAGE:685163"
/clone_11d="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"

BASE COUNT 85 a 88 c 90 g 53 t
ORIGIN
Query Match 100.0%; Score 9; DB 10; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 138 TTCGACAA 146
|||||
Cp 25 ttctgagaa 33

RESULT 195
LOCUS AA230352 316 bp mRNA EST 12-MAR-1997
DEFINITION mw99e01.r1 GuayWoodford Beier mouse kidney day 7 Mus musculus CDNA
clone IMAGE:663192 5', mRNA sequence.
ACCESSION AA230352
NID 91852668
VERSION AA230352.1 GI:1852668
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 316)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407450.

CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
NCI:409040
Seq primer: -28ml3 rev1 ET from Amersham.
Location/Qualifiers
1. 316
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: kidney; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
-5' adaptor sequence: 5' GATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTT 3' Library provided
by Lisa Guay-Woodford."
/db_xref="taxon:10090"
/map="4"
/clone="IMAGE:663192"
/clone_11d="GuayWoodford Beier mouse kidney day 7"
/tissue_type="kidney"
/dev_stage="juvenile (7 days old)"
/lab_host="SOLR (kanamycin resistant)"

BASE COUNT 117 a 47 c 46 g 106 t
ORIGIN
Query Match 100.0%; Score 9; DB 10; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.23e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 129 TTCGACAA 137
|||||
Cp 33 ttctcagaa 25

RESULT 196
LOCUS T12492 317 bp mRNA EST 28-NOV-1994
DEFINITION B079R Heart Homo sapiens CDNA clone B079, mRNA sequence.
ACCESSION T12492
NID 9597179
VERSION T12492.1 GI:597179
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 317)
 AUTHORS Liew, C.C., Hwang, D.M., Fung, Y.W., Laurensen, C., Cukerman, E., Tsui, S., and Lee, C.Y.
 TITLE A catalogue of genes in the cardiovascular system as identified by expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 10645-10649 (1994)
 MEDLINE 95024171
 COMMENT Other-ESTs: B079F
 Contact: Liew CC
 Department of Laboratory Medicine and Pathobiology
 University of Toronto
 Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
 Tel: 4169788758
 Fax: 4169785650
 Email: liewcc@utcc.utoronto.ca
 Seq primer: GACACGACCACTGGTAATG.
 Location/Qualifiers
 1..317
 /organism="Homo sapiens"
 /note="Vector: lambda gtl1. Site_1: EcoRI; Site_2: EcoRI"
 /db_xref="taxon:9606"
 /clone="B079"
 /clone_1lb="Heart"
 /lab_host="E.coli Y1090"

BASE COUNT 79 a 85 c 86 g 67 t

ORIGIN

Query Match 100.0%; Score 9; DB 8; Length 317;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 TTCGAGAA 31
 |||||||||
 25 ttcgagaa 33

RESULT 197
 LOCUS AA728030 317 bp mRNA EST 05-JAN-1998
 DEFINITION K458B554SL JH93SL-Bml4 (JHU) Brugia malayi cDNA clone B.malay1
 clone Y504 5' end, mRNA sequence.
 ACCESSION AA728030
 NID 92746945
 VERSION AA728030.1 GI:2746945
 KEYWORDS EST.
 SOURCE Brugia malayi.
 ORGANISM Brugia malayi.
 Eukaryota; Metazoa; Nematoda; Secernentea; Spirurida; Spirurida;
 Filarioidea; Onchocercidae; Brugia.
 1 (bases 1 to 317)
 Sabarinathan, R., Kannan, K., Rao, K.V.N., Ravl, V., Narayanan, R.B. and Jayaraman, K.
 Brugia malayi ESTs (Sabarinathan, R. et al.)
 TITLE Brugia malayi ESTs (Sabarinathan, R. et al.)
 JOURNAL Unpublished (1998)
 COMMENT On Nov 6, 1997 this sequence version replaced gi:931942.
 Contact: Kunthala Jayaraman
 Centre for Biotechnology
 Anna University
 Madras 600025, INDIA
 Tel: 091 44 2350299
 Fax: 091 44 2350299
 Email: cbte@brahma.ittm.ernet.in
 Seq primer: SL.
 Location/Qualifiers
 1..317
 /organism="Brugia malayi"
 /note="Vector: Uni-ZAP XR; Site_1: Xba I; Site_2: EcoR I;
 The cDNA library of the B.malay1 L4 stage was constructed
 at Johns Hopkins University. The presence of a conserved
 22 nucleotide splice leader at 5' end of large proportion
 of mature messages was exploited in the linked reverse
 transcriptase-polymerase chain reaction in the
 construction of this library. The library had unamplified

phage titre of 6.4x10E5 pfu per ml and 94% of the clones
 have inserts (mean length approx 700 bp).
 Acknowledgements: Nityakalyani Raghavan, Inca Ghosh,
 Ward Eisinger, Alan Scott Dept. of Molecular Microbiology
 and Immunology School of Public Health and Hygiene, Johns
 Hopkins University, 615, N.Wolfe street, Baltimore MD
 21205, USA. Tel 410 955 3442 Fax: 410 955 0105
 /db_xref="taxon:6279"
 /db_xref="taxon:6279"
 /clone="B.malay1 clone Y504"
 /clone_1lb="JH93SL-Bml4 (JHU)"
 /lab_host="Xl-1 Blue MRF"

BASE COUNT 88 a 60 c 66 g 103 t

ORIGIN

Query Match 100.0%; Score 9; DB 17; Length 317;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 196 TTCGAGAA 204
 |||||||||
 Cp 33 ttcgagaa 25

RESULT 198
 LOCUS AA626165 317 bp mRNA EST 15-OCT-1997
 DEFINITION zu90B10.S1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:745243
 3' mRNA sequence.
 ACCESSION AA626165
 NID 92538552
 VERSION AA626165.1 GI:2538552
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 317)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Lennon, G., Marra, M.,
 Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
 Thelning, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m3 fwd. ET from Amersham
 High quality sequence stop: 286.
 Location/Qualifiers
 1..317
 /organism="Homo sapiens"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5'
 TGTACCAATCTGAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /db_xref="GDB:5933120"
 /clone="IMAGE:745243"
 /clone_1lb="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"

BASE COUNT 86 a 69 c 83 g 79 t
 Query Match 100.0%; Score 9; DB 16; Length 317;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 152 TTCACGAA 160
 |||||
 Cp 33 ttctcagaa 25

RESULT 199
 LOCUS AA301326 319 bp mRNA EST 18-APR-1997
 DEFINITION EST14257 Testis tumor Homo sapiens CDNA 5' end similar to matrin 3,
 mRNA sequence.
 ACCESSION AA301326
 NID 91953659
 VERSION AA301326.1 GI:1953659
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
 Bait,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
 White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Well,C.,
 Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
 Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,T.A., Geophagen,N.S.,
 Glodde,A., Gnehm,C.L., Hanna,M.C., Hedlow,E., Hinkle,P.S.Jr.,
 Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
 He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
 Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
 Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
 Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
 Fraser,C.M. and Venter,J.C.
 TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1405130.
 Other ESTs: THCI68383
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3016699056
 Fax: 3016699423
 Email: akerlavet@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..319
 /organism="Homo sapiens"
 /note="Organ: testis; Vector: pBluescript SK-; Site_1:
 ECORI; Site_2: XhoI"
 /db_xref="ATCC (inhost):191679"
 /db_xref="taxon:9606"
 /map="12"
 /clone_lib="Testis tumor"
 /sex="male"
 /dev_stage="adult"
 /dev_stage="adult"
 BASE COUNT 123 a 49 c 87 g 57 t 3 others
 ORIGIN

Query Match 100.0%; Score 9; DB 11; Length 319;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 246 TTCACGAA 254
 |||||
 Oy 25 ttctcagaa 33

RESULT 200
 LOCUS AA929463 319 bp mRNA EST 23-APR-1998
 DEFINITION vt30h07.r1 Barstead mouse proximal colon MPLRB6 Mus musculus CDNA
 clone IMAGE:116463 5', mRNA sequence.
 ACCESSION AA929463
 NID 93078772
 VERSION AA929463.1 GI:3078772
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 1 (bases 1 to 319)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2153564.

CONTACT: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 THIS clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:630549
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 310.
 Location/Qualifiers
 1..319
 /organism="Mus musculus"
 /strain="FVB/N"
 /note="Vector: pT73D-PAC (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand CDNA
 was primed with a Not I - oligo(dT) primer [5',
 TCTTACGATCTCACTGAGAGCGCCGCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded CDNA was ligated to Eco RI adaptors
 [AATTCGAGCTCTG], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified pT73 vector.
 Library constructed by Bob Barstead.
 /db_xref="taxon:10090"
 /clone="IMAGE:116463"
 /clone_lib="Barstead mouse proximal colon MPLRB6"
 /dev_stage="7 day juvenile"
 /lab_host="DH10B"
 BASE COUNT 107 a 79 c 57 g 76 t
 ORIGIN

Query Match 100.0%; Score 9; DB 20; Length 319;
 Best Local Similarity 100.0%; Pred. No. 2.23e+03;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 203 TTCACGAA 211
 |||||
 Oy 25 ttctcagaa 33

Search completed: Thu Oct 28 00:29:41 1999

Thu Oct 28 08:43:28 1999

US-08-963-288-1.rst

Page 86

Job time : 228 secs.
